Cinnamic Acid-Specific Type III Polyketide Synthases of *Cajanus cajan* (L.) Millsp

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades

eines Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Su Zhang
aus Heilongjiang/ China

1. Referent: Professor Dr. Ludger Beerhues
2. Referent: apl. Professor Dr. Ulrich Engelhardt
eingereicht am: 24.06.2015
mündliche Prüfung (Disputation) am: 10.09.2015

Druckjahr 2015

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakulät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Tagungsbeiträge

Zhang S, Liu B, Beerhues L: Biosynthesis of cajanuslactone in pigeon pea. (Poster). DPhG Annual Meeting. Freiburg, Germany. October 9th-11th, 2013.

Acknowledgment

I would like to express my sincere appreciation, gratitude and deep respect to my supervisor Prof. Dr. *Ludger Beerhues* for giving me the opportunity to be a member of his research group and the valuable inspiring scientific discussions and his affectionate, friendly way of guidance and supervision. His continuous trust, unforgettable support and careful listening and comprehension made it possible to surpass the difficulties of work. Thanks for respecting and discussing my ideas and opinions, even when they appeared silly to me or not convincing to you. Working with you and in your workgroup taught me many lessons about science and life.

I owe much of the success of this work and most of the knowledge I expanded during this time to Dr. *Benye Liu*. I was so lucky that I could work with you in the same group in Germany. Thanks a lot for your help and guidance during all my life in Germany. Thanks for teaching me that before performing an experiment, it is so important to think about its outcome and methods and finally find a compromise between the resources and goals. For me, you are a great teacher who first gives his students all his experience to save their time and efforts and then the freedom to research freely. Thanks for your continuous support and respect even when I did mistakes and your brotherly advices and discussions.

A lot of thanks to Dr. *Till Beuerle* for his help with the chemical analysis, constructive discussions and valuable helpful advices during the chemical synthesis. I appreciate his welcoming approachable personality and kindness. Any time I could ask or discuss everything with him. Great thanks and appreciation also to Dr. *Rainer Lindigkeit* for his kind help with any software problems.

I would like to convey my thanks to my *past and present colleagues* and lab mates for the wonderful time we spent together, for their presence and assistance whenever needed. Thank you for being such an amazing group to work with. The variation of skills in our group helped me a lot during my laboratory work. I really appreciate the cooperative environment we worked in and I will miss the friendship of everyone in the lab. I would like to express my sincere gratitude to Mrs. *Ines Rahaus* and Mrs. *Doris Glindemann*. Thanks for your help with everyday working and your moral and affectionate support and continuous encouragement throughout the PhD work. Thanks for spreading optimism and fun in the stringent scientific life.

I extend my appreciation to my colleagues, *Dr. Mariam Gaid, Dr. Andreas Müller, Dr. Sahar Abdelaziz, Dr. Mohammed Khalil, Dr. Iman Abdel-Rahman, Tobias Fiesel, Dennis Reckwell, Islam El-Awaad, Mina Awadalah, Ebtesam Ali, Rabeia Ali, Marco Grull, and Mohamed Nagia* for their indispensable help and advice in everyday work. Thanks for sharing your experiences, successes and mistakes with me and the countless scientific discussions. Thanks for your kindness and support. I will always remember our happy and joyful conversations and laughter. My special thanks to Mrs. *Bettina Böttner* for the interesting conversations and her diligent work.

Many thanks to my *Professor Yujie Fu and colleagues* from the Key Laboratory of Forest Plant Ecology, Ministry of Education, Northeast Forestry University, Harbin, China for offering me this chance to study in Germany. I am greatly thankful to the *China Scholarship Council (CSC)* for funding this mission.

Words are not enough to express my hearty gratitude, sincere appreciation and great indebtedness to my *father*, *mother* and *brother* for their great care, trust, affection and love. Finally, I owe the success of this work to my *wife* who gave me the strength, encouragement and motivation to finish my PhD thesis.

……..To my wife,

and those whose contributions were most generous

Contents

Abbreviations

APS	Ammonium peroxydisulfate
ALDH	Aldehyde dehydrogenase
Bam	*Bacillus amyloli*
BIS	Biphenyl synthase
BLAST	Basic local alignment search tool
bp	Base pair
BSA	Bovine serum albumin
C	Degree Celsius
cDNA	Complementary deoxyribonucleic acid
Ci	Curie (unit of radioactivity)
CoA	Coenzyme A
2,4-D	2,4-dichlorophenoxyacetic acid
DAD	Diode array detector
DTT	1,4-dithiothreitol
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
EDTA	Ethylenediaminetetraacetic acid
EST	Expessed sequence tag
g	gramme
GC	Gas chromatography
h	Hour
Hin	*Haemophilus influenzae*
HPLC	High performance liquid chromatography
IPTG	Isopropyl--D-thiogalactopyranoside
K_m	Michaelis-Menten constant
K_{cat}	Turnover number
K_{cat} / K_m	Catalytic efficiency
LB	Luria broth
LS	Linsmaier and Skoog
m	Milli
M	Molar
MCS	Multiple cloning site
MS	Mass spectrometry
min	Minute
ml	Milliliter
mRNA	Messenger RNA
MSTFA	N-methyl-N-(trimethylsilyl)trifluoroacetamide
NAA	1-Naphthaleneacetic acid
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
NADP	Nicotinamide adenine dinucleotide phosphate (oxidized form)
Ni-NTA	Nickel-nitrilotriacetic acid
nm	Nanometer
OD	Optical density

ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PKS	Polyketide synthase
PMSF	Phenylmethylsulfonyl
RNA	Ribonucleic acid
rpm	Revolution per minute
RT	Reverse transcription
RT-PCR	Reverse transcription polymerase chain reaction
s	Second
SAM	S-adenosyl-L-methionine
SDS	Sodium dodecyl sulfate
TAE	Tris-acetate-EDTA
Taq	*Thermus aquaticus*
TEMED	N,N,N`,N`-tetramethylethylenediamine
T_m	Melting temperature (primer)
T_a	Annealing temperature
Tris	Tris (hydroxymethyl) aminomethane
UV	Ultraviolet

Amino acids

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

Nucleotides

A	Adenine
C	Cytosine
G	Guanine
T	Thymine
U	Uracil

I. Introduction

1. *Cajanus cajan*

Cajanus cajan (L.) Millsp. (Fig. 1), also known as pigeonpea, is one of the major grain legume crops grown in the semiarid tropics (SAT) extending between 30°N and 30°S. It is grown extensively in India and other countries of Southern Asia, Africa, and Latin America. It is an erect, branched, hairy shrub, 0.5-4 meters high. Leaves are oblong-lanceolate to oblanceolate with three leaflets. Flowers are yellow, in sparse peduncled racemes, about 1.5 cm long. Pods are hairy, 4-7 cm long, 1 cm wide, containing two to seven seeds (Fujita et al. 2004). *C.cajan* belongs to the genus *Cajanus* of the subtribe Cajaninae, tribe Phaseoleae of the subfamily Faboideae under the family Fabaceae. The genus *Cajanus* comprises 32 species, most of which are found in India and Australia although one is native to West Africa. *C.cajan* is the only cultivated food crop of the Cajaninae subtribe and has a diploid genome comprising 11 pairs of chromosomes (2n= 22) with a physical size estimated at about 0.853 pg (Greilhuber and Obermayer, 1988). In terms of global grain legume production, it ranks the sixth following soybean (*Glycine max* L.), common bean (*Phaseolus vulgaris* L.), chickpea (*Cicer arietinum* L.), broad bean (*Vicia faba* L.) and mungbean (*Vigna radiata* L. Wilczek) (Young et al., 2003).

Fig. 1. *Cajanus cajan* (L.) Millsp. (source: PickMeYard, 2010).

India is the world's largest *C. cajan* producer and accounts for over 77% of the total world production (Fig. 2). The current global annual production of *C. cajan* is valued at

approx. 4.68 m. tonnes (FAOSTAT 2013). The crop can be described as unique because it is a legume and a woody shrub. It has an inherent ability to withstand environmental stresses (especially drought) making it one of the most sought-after crops in plant introduction trials aimed at bringing new areas under cultivation (Okiror, 1986). It contributes to the C, N and P economy of the soil enhancing its performance even under marginal input (Fujita et al. 2004; Kumar Rao et al. 1987; Rego and Nageswara Rao, 2000). *C. cajan* is tolerant to low P supply and acid soils as well as having a high capacity for incorporation of external P into organic P (Fujita et al., 2004). Its critical requirement of P concentration for dry matter production is low compared to other major protein crops like soybean (*Glycine max*) (Adu et al., 1990).

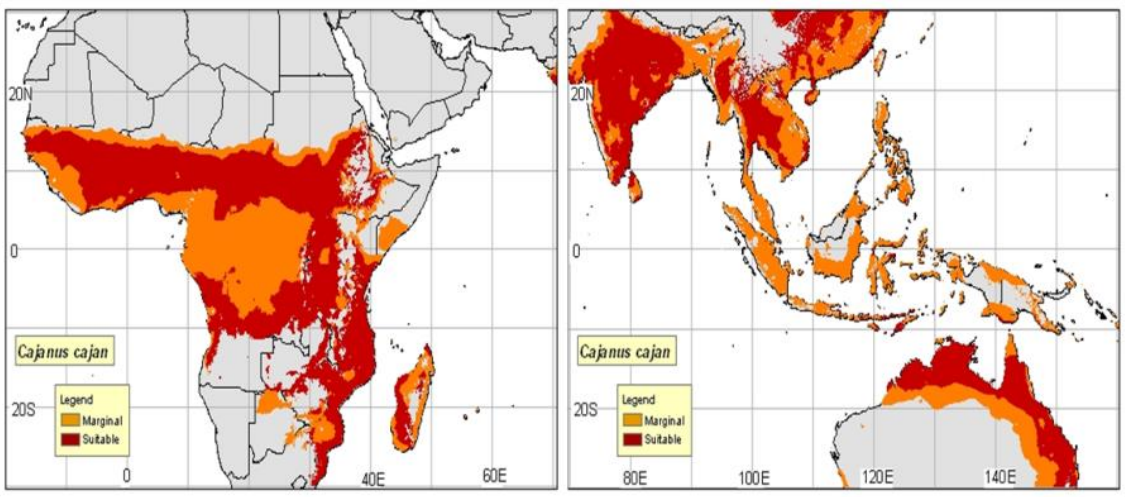


Fig. 2. The main producing areas of *C. cajan* in the world. (source: www.cwrdiversity.org/distribution-map/).

C. cajan plays an important role in food security, balanced diet and alleviation of poverty because it can be used in diverse ways as a source of food, feed, fodder (Rao et al., 2002), fuel wood, rearing *lac* insects (Zhenghong et al., 2001), hedges, windbreaks, soil conservation, green manuring and roofing. It is grown on about 5.25 million ha, yielding 3 million tons, and contributes to about 5% of total world production of pulses. It is also a major source of protein to about 20% of the world population (Thu et al., 2003) and is an abundant source of minerals and vitamins (Saxena et al., 2002). Its abundance in protein makes it an ideal supplement to traditional cereal-, banana- or tuber-based diets of resource poorfarmers who are generally protein-deficient. The perennial nature of *C. cajan* allows farmers to take multiple harvests with surpluses traded in both local and international markets.

C. cajan, being a forage crop, has been utilized as an important remedy for various ailments. The garo tribal community of Bangladesh utilizes it for the treatment of diabetes and as an energy stimulant (Pal et al., 2011). In Trinidad and Tobago the leaves of *C. cajan* are used in food poisoning, as colic and in constipation (Lans, 2007). In Chinese folk medicine, pigeon pea leaves are used to staunch blood, as an analgesic and to kill parasites. In some parts of Tamil Nadu, India, the leaf, seeds and young stems are used to cure gingivitis, stomatitis and as a toothbrush (Ganeshan, 2008). It is also an

important folk medicine in eastern Rajasthan as fresh juice or boiled leaves are given orally to nullify the effect of intoxication and as a laxative. Leaf paste is applied in oral ulcers and inflammations. Leaves and seeds are applied as poultice over the breast to induce lactation (Upadhyay et al., 2010). Amongst its many medicinal uses, most of the literature is focused on pigeonpea leaves, which demonstrated notable anti-inflammatory, anti-bacterial and abirritative properties (Fu et al., 2007). Especially in recent years, *C. cajan* leaves is of increasing importance and has been brought to market as a product of traditional Chinese medicine (TCM) for the therapy of ischemic necrosis of femoral head. It exhibits notable effects on the activation of blood circulation to dissipate blood stasis, reduce swelling, ease pain, and reinforce the kidney to strengthen the bone.

2. Phytochemistry of *C. cajan*

Plant secondary metabolite is a generic term used for more than 200,000 different substances.The plants form secondary metabolites e.g. for protection against pests, as colouring, scent, orattractants to pollinators and as the plant's own hormones (Wu and Chappell, 2008). It was believed that secondary metabolites were irrelevant for the human diet. However, secondary metabolites carry out a number of protective functions in the human body. These metabolites can boost the immune system, protect the body from free radicals, and kill pathogenic germs and much more (Makkar et al., 2007). In contrast to the primary metabolites (carbohydrates, fats, proteins, nucleic acids, vitamins and mineral nutrients) secondary metabolites are present in smaller amounts and do not have nutrient characteristics for human beings. They have aneffect on humans and health (Makkar et al., 2007). Therefore, in order to understand the mechanisms involved in these beneficial effects, a great deal of scientific efforts have been contributed to isolate and identify the active components of *C. cajan* during the last few decades. Extensive studies have been carried out regarding the chemistry of *C. cajan* and considerable progress has been achieved regarding its biological activities and medicinal applications.The phytochemistry of *C. cajan* is very complex; more than 380 compounds have been identified (Grover et al., 2002) representing different chemical classes. Some belong to primary metabolism, e.g. amino acids, fatty acids and steroids, while flavonoids, stilbenoids, stilbenecarboxylates, isocoumarins, anthraquinones, terpenoids, lignans and alkaloids represent secondary metabolites. The concentrations of these compounds depend on tissue type, age, variety, growth conditions (nutrition, humidity, light levels), harvest time and storage conditions (Liu et al., 2010). Different parts of *C. cajan* have been utilized for their biological activities since time immemorial and some of them have experimental grounds for their acceptance. Apart from their uses in folkloric medicines, there are several reports on the biological activities and pharmacological actions of *C. cajan* based on modern scientific investigations. A new natural isocoumarin, cajanuslactone A, has been isolated from the leaves of *C. cajan* which is a potential antibacterial agent against Gram-positive micro-organisms (Luo et al., 2008) The two stilbeneslongistylin A and longistylin C from leaves have been found to possess hypocholesterolemic effects (Luo et al., 2008). Anti-plasmodial activities have also been confirmed with betulinic acid isolated from roots and longistylin A and C obtained from leaves (Ezike et al., 2010). Pinostrobin, a substituted flavanone isolated from leaves possesses anti-inflammatory activity and inhibits sodium channel-activated

depolarization of mouse brain synaptoneurosomes (Nicholson et al., 2010). Four flavonoids, vitexin, apigenin, luteolin and orientin, isolated from the leaves were found to possess antioxidant activity (Zhang et al., 2010). Cajanol, an isoflovanone from the roots was found to possess anticancer activity (Luo et al., 2010). One important stilbenecarboxylate, cajaninstilbene acid, isolated from ethanolic extracts of leaves was found to possess significant antioxidant properties (Wu et al., 2009) as well as hypoglycemic, hypotriglyceridemic, hypocholesterolemic activities and potential action in the treatment of postmenopausal osteoporosis (Inman and Hoppe, 2002; Luo et al. 2008). Some protein fraction isolated from leaves also showed hepato-protective effects (Ahsan et al., 2009) and the presence of phenolics (tannins) and terpenoids impart anthelmintic activity (Singh et al., 2010). The major chemical constituents and their biological activities are summarized in Table 1.

Table 1. The main active constituents and pharmacological activities of *C. cajan*

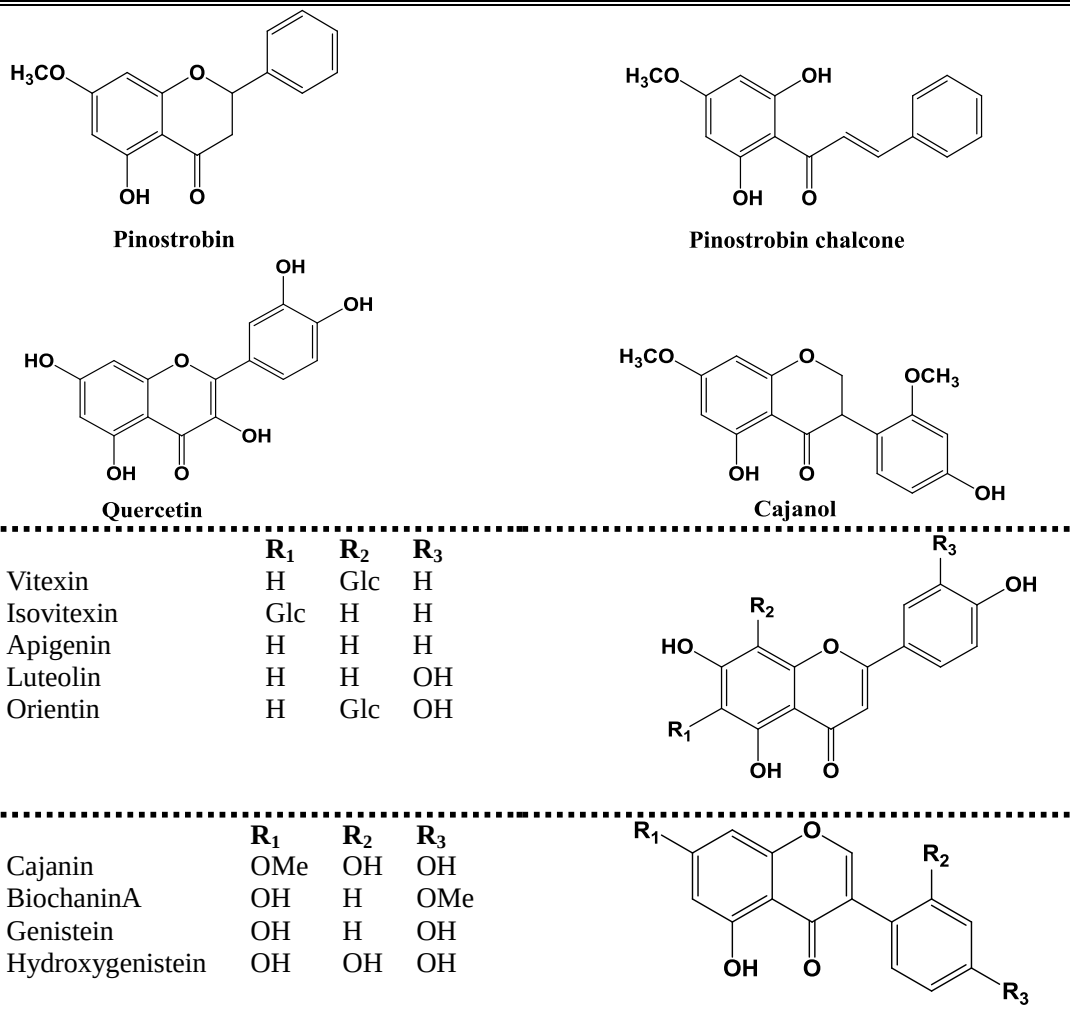
Class	Constituent	Pharmacological activity	Plant part
Flavonoid	Vitexin	Antioxidant	Leaf
	Isovitexin	Antioxidant	Leaf
	Apigenin	Antioxidant	Leaf and root
	Luteolin	Antioxidant	Leaf
	Orientin	Antiviral, antimicrobial, antioxidant	Leaf
	Cajanin	Hypocholesterolemic	Leaf
	Cajanol	Anticancer	Root
	Genistein	Antioxidant	Root
	Pinostrobin	Antispasmodic, anti-aromatase, anti-inflammatory	Leaf
	Pinostrobin chalcone	Antileishmanial, antioxidant, neuroprotective	Leaf
Stilbene	longistylin A	Antiplasmodial, antioxidant	Leaf
	longistylin C	Antiplasmodial, antitumoral, antioxidant	Leaf
Stilbenecarboxylate	Cajaninstilbene acid	Antibacterial, antioxidant, hypoglycemic, hypotriglyceridemic, hypocholesterolemic, treatment of postmenopausal osteoporosis	Leaf
Isocoumarin	Cajanuslactone A	Antibacterial	Leaf

Pharmacological investigations have indicated that *C. cajan* leaves produced most of the active compounds with the major bioactivities, such as antitumor, antibacterial, hypotriglycerimic, hypoglycemic, insecticidal, antiplasmodial, antioxidant activities, etc. (Ashidi et al., 2010). Chemical studies revealed that flavonoids, stilbenes, stilbenecarboxylates and isocoumarins represented the main active components found in *C. cajan* leaves (Ingham et al. 1976; Ohwaki et al., 1993).

2.1 Flavonoids

Flavonoids are a group of plant polyphenolic secondary metabolites showing a common three ring chemical structure (C_6-C_3-C_6). They are widely distributed in the plant kingdom and play an important role in the interaction between plants and their environment. The major classes of flavonoids are chalcones, flavones, flavonols, flavandiols, anthocyanins, and proanthocyanidins or condensed tannins. These compounds are widely distributed in different amounts, according to the plant species, organ, developmental stage and growth conditions. More than 6000 different flavonoids have been identified, and surely this number will increase (Ferrer et al., 2008). The different flavonoids have diverse biological functions, including protection against ultraviolet (UV) radiation and phytopathogens, signaling during nodulation, male fertility, auxin transport, as well as the coloration of flowers as a visual signal that attracts pollinators (Mol et al., 1998; Winkel-Shirley, 2002; Bradshaw and Schemske, 2003). Flavonoids are also responsible for the display of fall color in many plants, which may protect leaf cells from photooxidative damage, enhancing the efficiency of nutrient retrieval during senescence (Field et al., 2001). Flavonols are probably the most important flavonoids participating in stress responses; they are the most ancient and widespread flavonoids, having a wide range of potent physiological activities (Stafford, 1991; Pollastri and Tattini, 2011).

Studies of flavonoids production in *C. cajan* were rather sporadic. However, in recent years, in order to understand the mechanism involved in the beneficial effects of *C. cajan*, a great deal of scientific efforts have been contributed to isolate and identify the flavonoids in *C. cajan*. Chemical studies revealed that flavonoids are the major components found in *C. cajan* tissues, with the exception of the nonflavonoid stilbenecarboxylates, which are also important phenolics in *C. cajan* (Duker et al., 2004). To date, more than 20 flavonoids have been reported (Picerno et al., 2003) and the structures of the main flavonoid in *C. cajan* are summarized in Fig. 3. The major classes of flavonoids in *C. cajan* are flavones, isoflavons, chalcones, flavonols and flavanones. They are accumulated mainly in leaves and roots and these two parts exhibit the main pharmacological effects found in the whole plant.



	R_1	R_2	R_3
Vitexin	H	Glc	H
Isovitexin	Glc	H	H
Apigenin	H	H	H
Luteolin	H	H	OH
Orientin	H	Glc	OH

	R_1	R_2	R_3
Cajanin	OMe	OH	OH
BiochaninA	OH	H	OMe
Genistein	OH	H	OH
Hydroxygenistein	OH	OH	OH

Fig. 3. Structures of the main flavonoids isolated from *C. cajan*

2.2 Stilbenes

Stilbenes are natural phenolic compounds occurring in a number of plant families including Vitaceae, Dipterocarpaceae, Gnetaceae, Pinaceae, Fabaceae, Poaceae and Cyperaceae. Although polyphenolics display anenormous chemical diversity, stilbenes seem to constitute a rather restricted group of molecules. Most plant stilbenes are derivatives of the basic unit trans-resveratrol (3,5,4'-trihydroxy-*trans*-stilbene), although other structures are found in particular plant families. Examples of common plant stilbenes isolated from diverse families, including grape (Vitaceae), pine (Pinaceae), peanut (Fabaceae) and sorghum (Poaceae), are given in Fig. 4. Over the last 20 years, plant stilbenes have received considerable interest, due to their biological activities and possible pharmacological applications. Many roles have been ascribed to stilbenes, namely, as antimicrobial, deterrent, or repellent compounds in plants, protecting them from attacks by fungi, bacteria, nematodes, or herbivores (Jeandet et al., 2002; Chong et al., 2007). More recently, stilbenes (especially resveratrol and its derivatives) were acclaimed for their wondrous effects and wide range of purported healing and preventive

powers as cardioprotective, antitumor, neuroprotective, and antioxidant agents (Anekonda et al., 2006; Athar et al., 2007; Baur et al., 2006; King et al., 2006; Saiko et al., 2008; Jeandet et al., 2009).
Since resveratrol was postulated to be involved in the health benefits associated with a moderate consumption of red wine (Siemann et al., 1992), it is one of the most extensively studied natural products. Hundreds of studies have reported that resveratrol can prevent or slow the progression of a wide variety of illnesses, including cancer and cardiovascular diseases, as well as extend the lifespans of various organisms (Bauret al., 2006). The major role ascribed to stilbenes in a number of plant families is to act as phytoalexins (Langcake et al., 1977) and these stilbene phytoalexins are mainly derived from resveratrol or pinosylvin.

Stilbene	Occurence	R_3	R_5	R_3'	R_4'
Trans-resveratrol	Vitis, Arachis, Fallopia	OH	OH	H	OH
Trans-piceid	Vitis	OGlc	OH	H	OH
Pinosylvin	Pinus	OH	OH	H	H
Piceatannol	Picea	OH	OH	OH	OH
Pinosylvin monomethylether	Pinus, Alnus	OCH_3	OH	H	OH
Trans-pterostilbene	Vitis, Vaccinium	OCH_3	OCH_3	H	OH
Astringin	Picea	OGlc	OH	OH	OH
Rhapontin	Rheum	OGlc	OH	OH	OCH_3

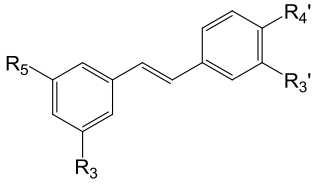


Fig. 4. Stuctures of common plant stilbenes. Examples of plant genera where these stilbenes are known to occur are also mentioned. OGlc: O-ß-D-glucopyranoside.

Although only few stilbene species were identified in *C. cajan*, chemical studies revealed that stilbenes represented the main active components found in *C. cajan* leaves (Ohwaki et al., 1993). This is because of two important stilbenes, longistylin A and longistylin C, which are found only in *C. cajan* and are unique secondary metabolites derived from pinosylvin (Fig. 5). As representative stilbenes in pigeon pea leaves, longistylin A and longistylin C showed a significant high *in vitro* activity against the chloroquine-sensitive *Plasmodium falciparum* strain 3D7 (Luo et al., 2008). They have also been found to possess hypocholesterolemic effects (Ezike et al., 2010). Longistyline C also exhibited excellent inhibitory activity on various cancer cells. Even more interesting is that longistylin A and longistylin C are considered to be associated with the biosynthesis of stilbenecarboxylates in *C. cajan* (Cooksey et al., 1982).

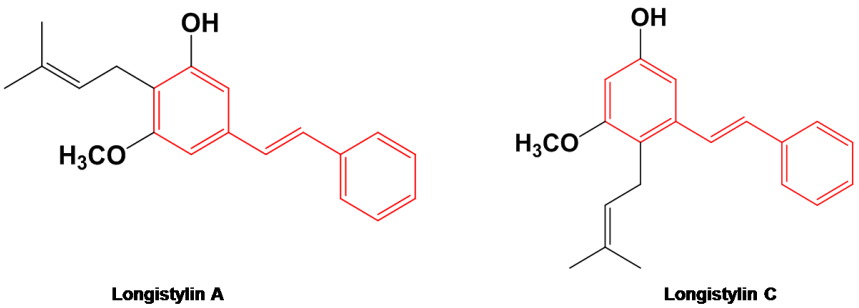


Fig. 5. Stuctures of two stilbenes from *C. cajan*.

2.3 Isocoumarin

α-Pyranone (2*H*-pyran-2-one), a six-membered oxygen heterocycle (Fig. 6) represents an important class of naturally occurring unsaturated δ lactone. Its benzo derivatives, e.g. isocoumarins (Fig. 6), are structurally related to another naturally occurring lactone called coumarin but with an inverted lactone ring as indicated by their name (Barry et al., 1964; Napolitano et al., 1997).

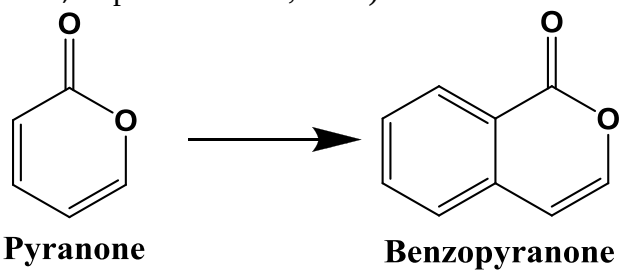


Fig. 6. Structures of pyranone and benzopyranone.

The prevalence of isocoumarins in numerous natural products that exhibit a wide range of biological activities has generated acontinued and enormous interest among synthetic and medicinal chemists. The isocoumarin framework represents one of the privileged structures for the development of natural product-inspired compounds of potential biological interest. Like coumarins, this class of aromatic lactones is of considerable interest because of their natural occurrence (Wijeratne et al., 2006) and a wide range of pharmacological activities such as antifungal (Nozawa et al., 1981), anti-inflammatory (Furuta et al., 1986), antimicrobial (Matsuda et al., 1999), phytotoxic (Sato et al., 1981), cytotoxic (Whyte et al., 1996), and other effects. For example, thunberginols, isolated from *Hydrangea Dulcis Folium*, the fermented and dried leaves of *Hydrangea macrophylla* Seringe var. thumbergii Makino showed antidiabetic properties (Zhang et al., 2007). A naphthopyranone dimer, named planifolin was isolated from the capitula of *Paepalanthus planifolius* Koern and showed cytotoxic as well as mutagenic activities (Varanda et al., 2006). Recently, new isocoumarins isolated from *Microdochiumbolleyi*, an endophytic fungus from *Fagonia cretica* (a herbaceousplant of the semiarid coastal regions of Gomera) have shown good antifungal, antibacterial, and antialgal activities against *Microbotryum violaceum*, *Escherichia coli*, *Bacillus megaterium*, and *Chlorella fusca* (Zhang et al., 2008). The isocoumarin NM-3 (Oikawa et al., 1997), an analogue of the natural product cytogenin, inhibited the growth of human endothelial cells in culture and tumor angiogenesis in human tumor xenograft models. Significant reductions in

mean tumor volume were observed in animal models when NM-3 was administered in combination with other chemotherapeutic agents or radiation, beyond those observed with chemotherapy or radiotherapy alone. Subsequently after discovery of NM-3, angiogenesis has become anactive area of pharmaceutical research. Some examples of naturally occurring isocoumarins and related derivatives are presented in Fig.7. Isocoumarins are also useful intermediates in the synthesis ofa variety of important hetero- and carbocyclic compounds, including isocarbostyrils, isoquinolines, isochromenes, and various aromatic compounds.

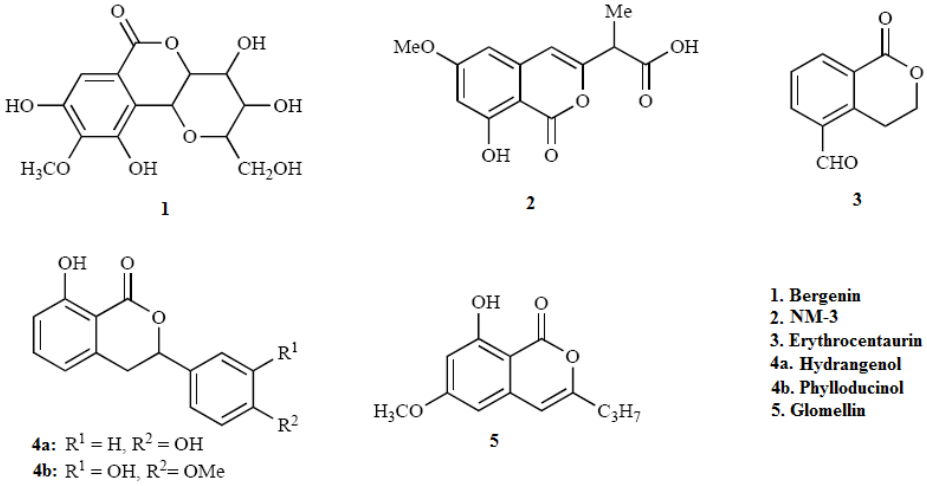


Fig. 7. Some naturally occurring isocoumarins.

Two new natural isocoumarins, cajanuslactone A and cajanuslactone C (Fig. 8), have been isolated from the leaves of *C. cajan* which are unique secondary metabolites found only in *C. cajan.* Cajanuslactone A is a potential antibacterial agent against Gram-positive micro-organisms and exhibited excellent antioxidant activity (Liu et al., 2010; Luo et al., 2008).

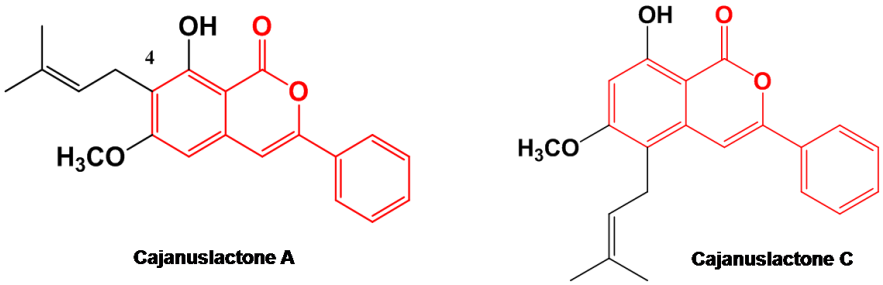


Fig. 8. Structures of two new natural isocoumarins, cajanuslactone A and cajanuslactone C from the leaves of *C. cajan.*

2.4 Stilbenecarboxylate

Stilbenecarboxylates (also known as stilbenecarboxylic acids or stilbene acids, Fig. 9) are a kind of unique secondary metabolites which are rare in plants, only few such compounds appear to be known. Like stilbenes, the rare natural products have been

isolated from a small but diverse collection of plants. Lunularic acid **(**Fig. 10) has been isolated from hortensia (*Hydrangea macrophylla*), the liverworts *Lunularia cruciata* (Eckermann et al., 2003), *Riella* spp., *Marchantia polymorpha*, *Blasia pusilla*, and *Riccia* spp. (Asakawa et al., 2004), and celery (*Apium graveolens*) (Zhou et al., 2009). It has growth inhibitory and dormancy-inducing effects for lower plants (Yoshikawa et al., 2002). It also shows fungicidal, algicidal and antihyaluronidase activities (Asakawa et al., 2007). Hydrangeic acid **(**Fig. 10) is a stilbenecarboxylate constituent of hortensia (*Hydrangea macrophylla*). Hydrangeic acid possesses anti-diabetic activity and lowers blood glucose, triglyceride and free fatty acid levels (Zhang et al., 2009). Pinosylvic acid, (Fig. 10) is another stilbenecarboxylate found in climbing skullcap (*Scutellaria scandens*). The leaves of this plant are traditionally used to treat wounds and swelling by insects (Shang et al., 2010).The 4-*O*-β-D-glucopyranoside derivative of pinosylvic acid, called gaylussacin (Fig. 10), is found in blackhuckleberry (*Gaylussacia baccata*), dangleberry (*Gaylussacia frondosa*) and climbing skullcap (*Scutellaria scandens*) (Malikov et al., 2002).

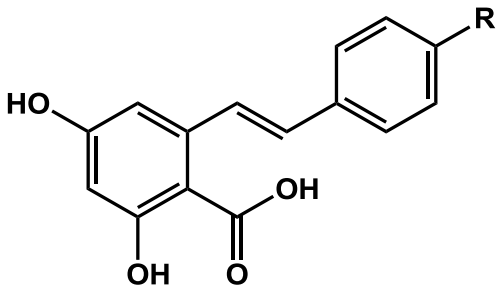


Fig. 9. Stilbenecarboxylate skeleton.

In *C. cajan*, three stilbenecarboxylates, cajaninstilbene acid, isocajaninstilbene acid and 4-*O*-methylpinosylvic acid (Fig.10) were isolated from the leaves long time ago and all of them are unique secondary metabolites found only in *C. cajan* (Cooksey et al., 1982; Yannai, 2004). Recent chemical studies revealed that stilbenecarboxylates represented the main active components found in *C. cajan* leaves. Most of studies on pharmacological activity of *C. cajan* have focused on the stilbenecarboxylates. They are becoming the most prevalent compounds in *C. cajan* in recent years. At present, in China, *C. cajan* leaves have been exploited in traditional Chinese medicine (TCM) for the therapy of ischemic necrosis of the femoral head, and pharmacological investigations found cajaninstilbene acid (CSA), a stilbenecarboxylate, is considered to be the most important active ingredient of this medicine (Fu et al., 2007). As representative stilbenecarboxylate in *C. cajan*, cajaninstilbene acid had estrogenic, hypoglycemic, hypotriglyceridemic, and hypocholesterolemic activities as well as exhibited excellent inhibitory activity on various cancer cells (Luo et al. 2008; Wu et al., 2009).

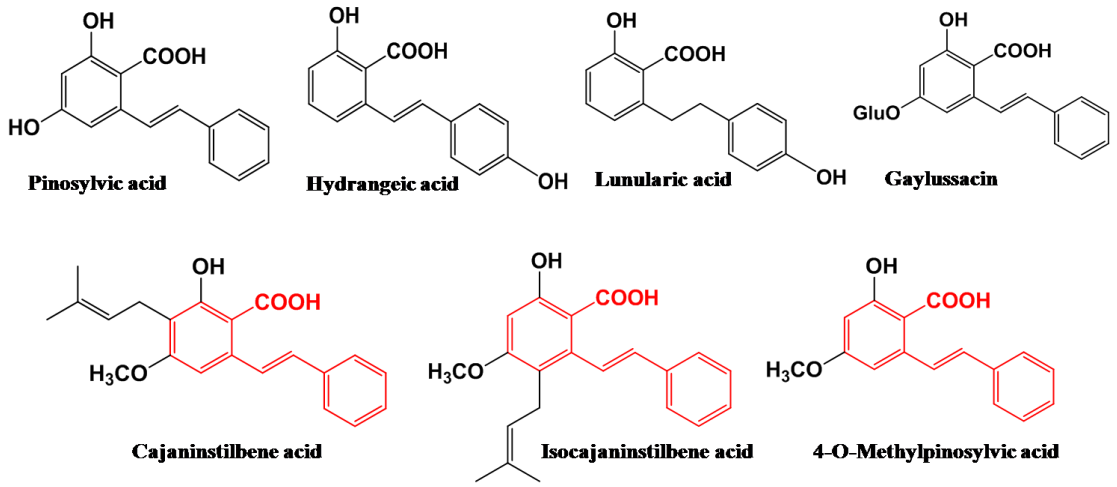


Fig. 10. Structures of some stilbenecarboxylates discussed in the text.

3. The biosynthesis of main secondary metabolites in *C. cajan*

3.1 Plant type III polyketide synthase (PKS) enzymes

The polyketide synthases (PKSs) are a group of enzymes that catalyze the condensation of CoA-esters of acetic acid and other acids to give polyketide compounds. They are classified according to their architectural configurations as type I, II and III (Hopwood and Herman, 1990; Staunton and Weissman, 2001; Fischbach and Walsh, 2006). The type I PKS describes a system of one or more multifunctional proteins that contain a different active site for each enzyme-catalyzed reaction in polyketide carbon chain assembly and modification. They are organized into modules, containing at least acyltransferase (AT), acyl carrier protein (ACP) and β-keto acyl synthase (β-KS) activities. Type I PKSs are sub-grouped as iterative or modular, usually present in fungal or bacterial systems, respectively (Moore and Hopke, 2001; Moss et al., 2004). The type II PKS is a system of individual enzymes that carry a single set of iteratively acting activities, and a minimal set consists of two ketosynthase units (α- and β-KS) and an ACP, which serves as an anchor for the growing polyketide chain. Additional PKS subunits such as ketoreductases, cyclases or aromatases define the folding pattern of the polyketo intermediate and further post-PKS modifications, such as oxidations, reductions or glycosylations are added to the polyketide (Rix et al., 2002). The only known group of organisms that employs type II PKS systems for polyketide biosynthesis is soil-borne and marine Gram-positive actinomycetes. The type III PKS is present in bacteria, plants and fungi (Austin and Noel, 2003; Funa et al., 2007). They are essentially condensing enzymes that lack ACP and act directly on acyl-CoA substrates.

A number of type III polyketide synthases (PKSs) have been found in plants to generate an amazingly diverse array of secondary metabolites by catalyzing the sequential condensation of acetyl units derived from the malonyl-CoA thioester into a growing polyketide chain (Austin and Noel, 2003). This reaction sequence mirrors the biosynthetic pathway of the fatty acid synthase enzymes of primary metabolism, from which the PKSs likely evolved via gene duplication events and subsequent refunctionalization of the gene duplicates. The PKSs in plants constitute the type III PKS

superfamily of enzymes, also called superfamily of CHS-like enzymes (Schröder, 2000; Austin and Noel, 2003). It is known that plant PKSs share 44-95% amino acid sequence identity andutilize a variety of different starter substrates ranging from aliphatic-CoA to aromatic-CoA substrates, from small (acetyl-CoA) to bulky (*p*-coumaroyl-CoA) substrates or from polar (malonyl-CoA) to nonpolar (isovaleroyl-CoA) substrates, giving the plants an extraordinarily functional and thereby chemical diversification (Flores-Sanchez and Verpoorte, 2009). This remarkable variety is further increased by divergences in the number of condensation reactions (polyketide chain elongation) and the type of the intramolecular cyclization reaction used (Schröder, 2000). Based on the mechanism of the cyclization they are classified as CHS-, STS- and CTAS-type (Table 2).

Table 2. Examples of type III polyketide synthases, preferred substrates and reaction products (Flores-Sanchez and Verpoorte, 2009).

Enzyme	Substrates (stater, extender, no. condensations)	Product	References
Plant:			
None cyclization reaction			
Benzalacetone synthase (BAS),	*p*-coumaroyl-CoA, Malonyl-CoA (1X)	Benzalacetone	Borejsza-Wysocki and Hrazdina, 1996; Abe *et al*., 2001; Zheng and Hrazdina, 2008
	Feruloyl-CoA, Malonyl-CoA (1X)	Methoxy-benzalacetone (**12**)	
One cyclization reaction			
Benzalacetone synthase (BAS),	*N*-methylanthraniloyl-CoA (or *N*-methylanthraniloyl-CoA (or anthraniloyl-CoA), Malonyl-CoA (or methyl-malonyl-CoA) (1X)	4-hydroxy-2(1*H*)quinolones (**3**)	Abe *et al*., 2006a
CTAS type			
C-methylchalcone synthase (PstrCHS2)	Diketide-CoA, Methyl-malonyl-CoA (1X)	Methyl-pyrone	Schröder *et al*., 1998
Styrylpyrone synthase (SPS) or Bisnoryangonin synthase	*p*-coumaroyl-CoA, Malonyl-CoA (2X)	Bisnoryangonin	Beckert *et al*., 1997; Herderich *et al*., 1997; Schröder Group
(BNS)	Caffeoyl-CoA, Malonyl-CoA (2X)	Hispidin	
2-pyrone synthase (2-PS)	Acetyl-CoA, Malonyl-CoA (2X)	Triacetic acid lactone (TAL)	Eckermann *et al*., 1998
p-Coumaroyltriacetic acid synthase (CTAS)	*p*-coumaroyl-CoA, Malonyl-CoA (3X)	*p*-coumaroyltriacetic acid lactone	Akiyama *et al*., 1999
CHS type			
Chalcone synthase (CHS)	*p*-coumaroyl-CoA, Malonyl-CoA (3X)	Naringenin chalcone	Whitehead and Dixon, 1983; Ferrer *et al*., 1999
Phlorisovalerophenone synthase (VPS)	Isovaleroyl-CoA, Malonyl-CoA (3)	Phlorisovalerophenone	Paniego *et al*., 1999; Okada and Ito, 2001
Isobutyrophenone synthase (BUS)	Isobutyryl-CoA, Malonyl-CoA (3X)	Phlorisobutyrophenone	Klingauf *et al*., 2005
Benzophenone synthase (BPS),	*m*-hydroxybenzoyl-CoA, Malonyl-CoA (3X)	2,3',4,6-tetrahydroxybenzophenone	Beerhues, 1996
	Benzoyl-CoA, Malonyl-CoA (3X)	2,4,6-trihydroxybenzophenone	Liu *et al*., 2003
Acridone synthase (ACS)	*N*-methylanthraniloyl-CoA, Malonyl-CoA (3X)	1,3-dihydroxy-*N*-methylacridone	Junghanns *et al*., 1998; Springo *et al*., 2000
Homoeriodictyol/ eriodictyol synthase (HEDS or HvCHS)	Feruloyl-CoA, Malonyl-CoA (3X)	Homoeriodictyol	Christensen *et al*., 1998
	Caffeoyl-CoA, Malonyl-CoA(3X)	Eriodictyol	
STS type			
Stilbene synthase (STS)	*p*-coumaroyl-CoA, Malonyl-CoA (3X)	Resveratrol	Schöppner and Kindl, 1984; Austin *et al*., 2004a
Pinosylvin synthase	Cinnamoyl-CoA, Malonyl-CoA (3X)	Pinosylvin	Raiber *et al*., 1995; Schanz et al., 1992; Fliegmann *et al*., 1992
Bibenzyl synthase (BBS)	Dihydro-*m*-coumaroyl-CoA, Malonyl-CoA (3X)	3,3',5-trihydroxybibenzyl	Reinecke and Kindl, 1994; Preisig-Müller *et al*., 1995
Biphenyl synthase (BIS)	Benzoyl-CoA, Malonyl-CoA (3X)	3,5-dihydroxybiphenyl	Liu *et al*., 2007
Stilbenecarboxylate synthase (STCS)	Dihydro-*p*-coumaroyl-CoA, Malonyl-CoA (3X)	5-hydroxylunularic acid	Eckermann *et al*., 2003; Schröder Group
More than 2 cyclization reactions			
Miscellaneous type			
Pentaketide chromone synthase (PCS)	Acetyl-CoA, Malonyl-CoA (4X)	5,7-dihydroxy-2-methylchromone	Abe *et al*., 2005a
Hexaketide synthase (HKS)	Acetyl-CoA, Malonyl-CoA (5 X)	6-(2',4'-dihydroxy-6'-methyl-phenyl)-4-hydroxy-2-pyrone	Springob *et al*., 2007; Jindaprasert *et al*., 2008
Aloesone synthase (ALS)	Acetyl-CoA, Malonyl-CoA (6X)	Aloesone	Abe *et al*., 2004a
Octaketide synthase (OKS)	Acetyl-CoA, Malonyl-CoA (7X)	SEK4 (**25**) and SEK4b (octaketides)	Abe *et al*., 2005b

In the CHS type, the intramolecular cyclization from C6 to C1 is called Claisen condensation; this mechanism for the carbon-carbon bond formation is not only used for the biosynthesis of polyketides, but also for fatty acids (Heath and Rock, 2002). In the STS type, the cyclization is from C2 to C7, with an additional decarboxylative loss of the C1 as CO_2, this reaction is an aldol-type condensation. In the CTAS type, there is a heterocyclic lactone formation between oxygen from C5 to C1, called lactonization. Thus, examples of type III PKSs that participate in the biosynthesis of secondary metabolites are chalcone synthase (CHS), stilbene synthase (STS), *p*-coumaroyltriacetic acid synthase (CTAS) and stilbenecarboxylate synthase (STCS) (Fig. 11). These enzymes are specific plant type III PKS that are involved in the formation of the mentioned natural compounds as components of secondary metabolism in *C. cajan*.

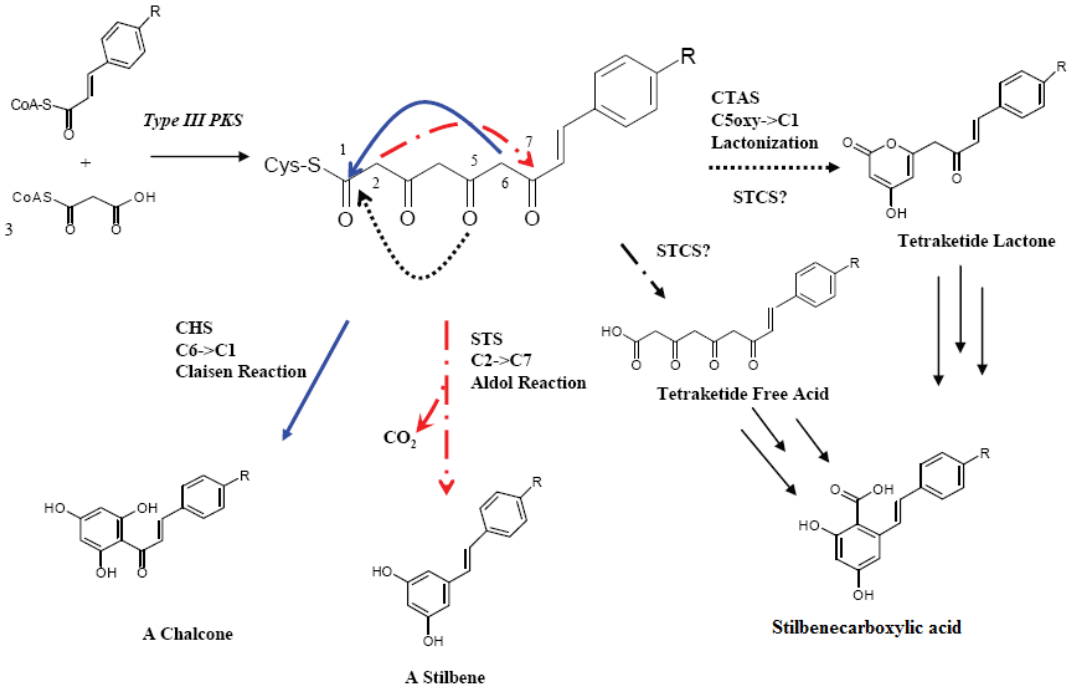


Fig. 11. Types of cyclization used by plant PKSs. R, OH or H (modified from Austin et al., 2004a).

3.2 Flavonoid biosynthesis

Flavonoids are synthesized along the general phenylpropanoid pathway by the activity of a cytosolic multienzyme complex, known also as flavonoid metabolon, loosely associated to the cytoplasmic surface of the endoplasmic reticulum. In particular, some of these enzymes belong to the cytochrome-P450 family and possess the ability to bind to membranes (Stafford et al., 1974; Winkel et al., 2004). On the other hand, some of the enzymes involved in the biosynthetic pathway are loosely associated with membranes of different organelles, such as vacuole (Kuhn et al., 2011; Ono et al., 2006; Saslowsky et al., 2005; Toda et al., 2012), plastids and nucleus (Wang et al., 2010; Tian et al., 2008; Wang et al., 2011). Much effort has been made in elucidating the biosynthetic pathways of flavonoids from a genetic perspective. Mutants affecting flavonoid synthesis were

isolated in a range of plant species. Maize (*Zeamays*), snapdragon (*Antirrhinummajus*), and petunia (*Petuniahybrida*) were established as the first major experimental models in this system, leading to the isolation of many structural and regulatory flavonoid genes (Holton et al., 1993; Mol et al., 1998). Although the flavonoid pathway has been extensively studied in several plants, there is no data on the biosynthesis of flavonoids in *C. cajan.*

The flavonoid biosynthetic pathway has largely been characterized (Austin and Noel, 2002) (Fig.12). Although chalcone synthase catalyzes the first committed step of the multi-branched flavonoid pathway, CHS is one of several enzymes using products of the upstream general phenylpropanoid pathway, which synthesizes *p*-coumaroyl-CoA from L-phenylalanine in three enzymatic steps (Fig. 12). Aside from flavonoid biosynthesis, intermediates and products of the general phenylpropanoid pathway are diverted to synthesize benzoic acid, salicylic acid, coumarins, and the monolignol precursors of lignin, an important structural and defensive polymer (Weisshaar et al., 1998; Dixon et al., 1995**).** The precursors are phenylalanine from the shikimate pathway and malonyl-CoA, which is synthesized by carboxylation of acetyl-CoA, a central intermediate in the Krebs tricarboxylic acid cycle (TCA cycle). Phenylalanine is converted into *trans*-cinnamic acid by a phenylalanine ammonia lyase (PAL); this cinnamic acid is hydroxylated by a cinnamate 4-hydroxylase (C4H) to *p*-coumaric acid and a CoA thiol ester is added by a 4-coumarate:CoA ligase (4CL). One molecule of *p*-coumaroyl-CoA and three molecules of malonyl-CoA are condensed by a chalcone synthase (CHS), a PKS yielding naringenin chalcone. The naringenin chalcone is subsequently isomerized by the enzyme chalcone isomerase (CHI) to naringenin, a flavanone. This naringenin is the common substrate for the biosynthesis of flavones and flavonols. Hydroxy substitution to ring C at position 3 by a flavanone 3-hydrolase (F3H) and to ring B at position 3' by a flavonoid 3'-hydrolase (F3'H) occur in naringenin. F3H is a 2-oxoglutarate-dependent dioxygenase (2OGD) and F3'H is a cytochrome P450. Subsequently, in the ring C at positions 2 and 3 a double bond is formed by a flavonol synthase (FLS), or flavone synthase (FNS). FLS is a 2ODG and for FNS two distinct activities have been characterized that convert flavanones to flavones. In most plants FNS is a P450 enzyme (FNSII), but in species of Apiaceae FNS is a 2ODG (FNSI). Modification reactions as glycosylation by UDP-glucose: flavonoid 3-*O*-glucosyl transferase (UFGT), methylation by a SAM-methyltransferase (OMT) and prenylation by prenyltransferases are added to the flavone and flavonol. Meanwhile, there are alternative routes for isoflavones which are formed by isoflavone synthase (IFS). Naringenin may also be directly hydroxylated by F3'H or F3'5'H to deliver eriodictyol and pentahydroxy flavanone, respectively, which are again hydroxylated to dihydroquercetin and dihydromyricetin. The three dihydroflavonols thus synthesized are then converted to anthocyanidins (coloured but unstable pigments) by two reactions catalysed by dihydroflavonol reductase (DFR) and anthocyanidin synthase (ANS). Modification reactions as glycosylation, methylation and prenylation by prenyltransferases are added in the last common step for the production of some specific compounds (Harborne et al., 2000; Winkel et al., 2001; Holton et al., 1995).

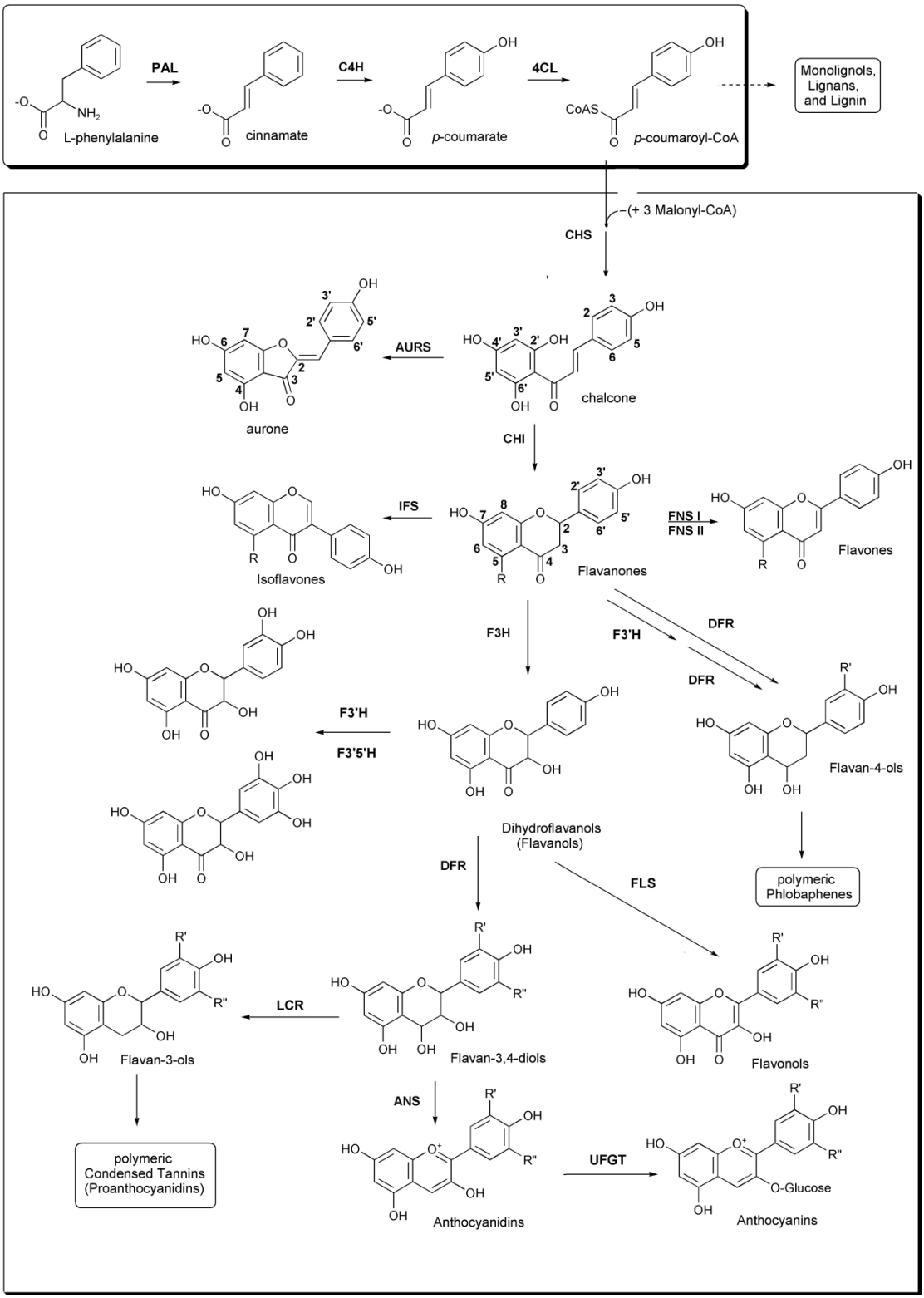


Fig. 12. Simplified diagram of the flavonoid biosynthetic pathway (Austin and Noel, 2002)

3.3 Stilbene biosynthesis

The plant stilbene biosynthetic pathway is closely related to the biosynthetic pathway of flavonoids, both are derived from the general phenylpropanoid pathway (Fig. 13). All higher plants seem to be able to synthesize malonyl-CoA and CoA-esters of cinnammic acid derivatives, but only few plant species are able to produce stilbenes. The first enzymes of the phenylpropanoid pathway have been described above, such as phenylalanine ammonia lyase (PAL), cinnamate-4-hydroxylase (C4H) and 4-coumarate: CoA ligase (4CL) (Ehlting et al., 2006; Ferreret al., 1999; MacDonald et al., 2007). Stilbenes are formed on the phenylalanine/ polymalonate route, the last step of this biosynthesis pathway being catalyzed by stilbenesynthase (STS) (Fig. 13). STS belongs to the type III of the polyketide synthase enzyme superfamily, a class of enzymes which carry out iterative condensation reactions with malonyl-CoA (Austin et al., 2003; Yu et al., 2008). STS produces the simple stilbene phytoalexins in one enzymatic reaction with starter coenzyme A-esters of cinnamic acid derivatives (*p*-coumaroyl-CoA in the case of resveratrol or cinnamoyl-CoA in the case of pinosylvin) and three malonyl-CoA units. Although it was suggested that STS has a monomeric structure, the homodimeric nature of STS comprising two 40-45 kDa subunits (M_r ca. 90 kDa) was later unambiguously demonstrated (Tropfet al., 1995). STS appears to have evolved from chalcone synthase (CHS) during land plant evolution and the enzymes share more than 65% amino acid homology. Unlike the bacterial PKSs, which are encoded by large gene clusters (Kao et al., 1994; Pfeifer et al., 2001), both STS and CHS functionas unimodular PKSs with a single active site, forming relatively small homodimers. Both CHS and STS use the same substrates and synthesize the same linear tetraketide intermediate (Fig. 12). However, while CHS cyclizes this intermediate by an intramolecular Claisen condensation, STS uses a different cyclization mechanism, involving an aldol condensation, accompanied by an additional decarboxylation leading to the loss of one carbon as CO_2. Early comparisons of CHS and STS proteins failed to identify consensus sequences that could explain the basis for these different cyclization mechanisms (Tropf et al., 1994). Homology modeling of STS structure based on CHS crystal structure did not predict any significant topological or chemical differences. Resolution of the crystal structure of STS from *P. sylvestris* and mutagenic conversion of alfalfa CHS into a functional STS showed that the STS aldol type of cyclization is mainly due to electronic effects rather than steric factors of the active site of the enzyme.The mutagenic conversion of alfalfa CHS into a functional STS involved eight mutations, which do not correspond to conserved residues in the STS enzymes from peanut or grapevine, confirming the lack of a universal STS consensus sequence (Austin et al., 2004).

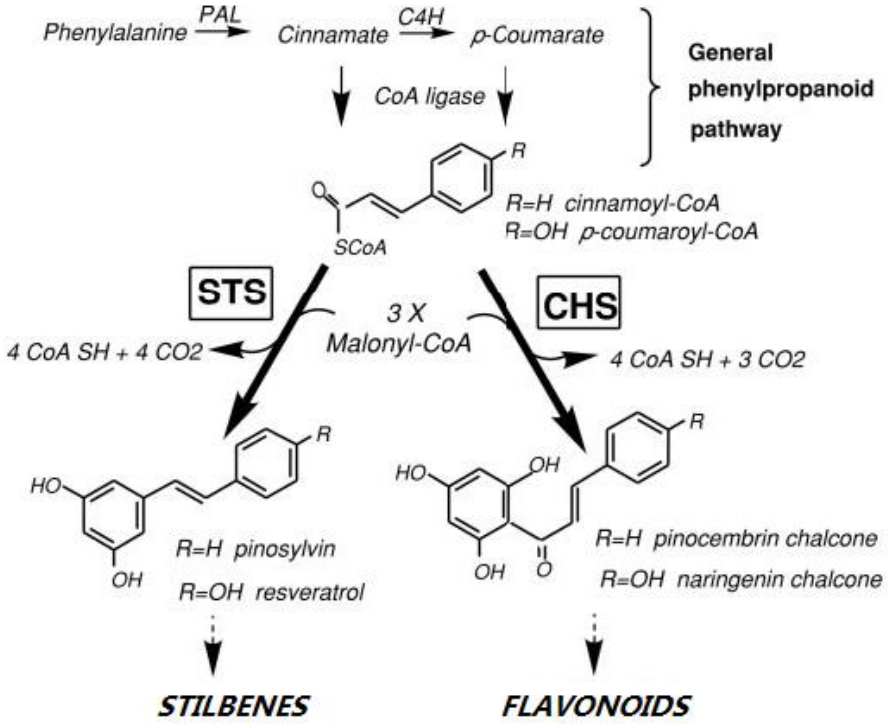


Fig. 13. Pathways for stilbene and flavonoid biosynthesis (Kodan et al. 2001).

STSs are classified into *p*-coumaroyl-CoA-specific type such as resveratrol synthase or cinnamoyl-CoA-specific type such as pinosylvin synthase, depending on their preferred starter molecule. The former type occurs mostly in angiosperms such as peanut (Schöppner et al., 1984), grapevine (Melchior et al., 1991) and sorghum (Yu et al., 2005), whereas the latter type is typical for gymnosperms and has been reported in several pine species such as *P. sylvestris* (Fliegmann et al., 1992), *P. strobes* (Raiber et al., 1995) and *P. densiflora* (Kodan et al., 2002). However, STS enzymes may accept different cinnammic acid derivatives as substrates and a single enzyme may be responsible for the biosynthesis of different stilbenes, depending on the starter molecule. STS was first purified from cell suspension cultures of *A. hypogaea* (Schoeppner et al., 1984). In *Vitis vinifera*, one of the plant species in which the highest amounts of resveratrol are naturally found, genome sequencing has revealed a large expansion of STS genes (43 genes identified with 20 of these previously beingshown to be expressed), suggesting the great importance of stilbene metabolism for this species (Jaillon et al., 2007). STS is encoded by amultigene family mainly comprising the resveratrol-forming STS genes from grapevine (*pSV21, pSV25, pSV696, pSV368* and *StSy*), (*Vst1*, *Vst2*, and *Vst3*) (Wiese et al., 1994), and the pinosylvin-forming STS genes from pine (*PST-1*, *PST-2*, *PST-3*, *PST-4*, and *PST-5*) (Preisig et al., 1999). Three novel STS genes (*pdsts1*, *pdsts2*, and *pdsts3*) have been isolated from the roots of *Pinus densiflora* (Kodan et al., 2001), together with a STS gene from *Vitis riparia* cv. Gloirede Montpellier (Goodwin et al., 2000). At present, there is only one STS gene described in a monocotyledonous plant, the *Sb STS1* gene isolated from sorghum. However, there is no any data on the biosynthesis of stilbenes in *C. cajan*. Some common modifications of plant stilbenes are shown in Fig. 14.

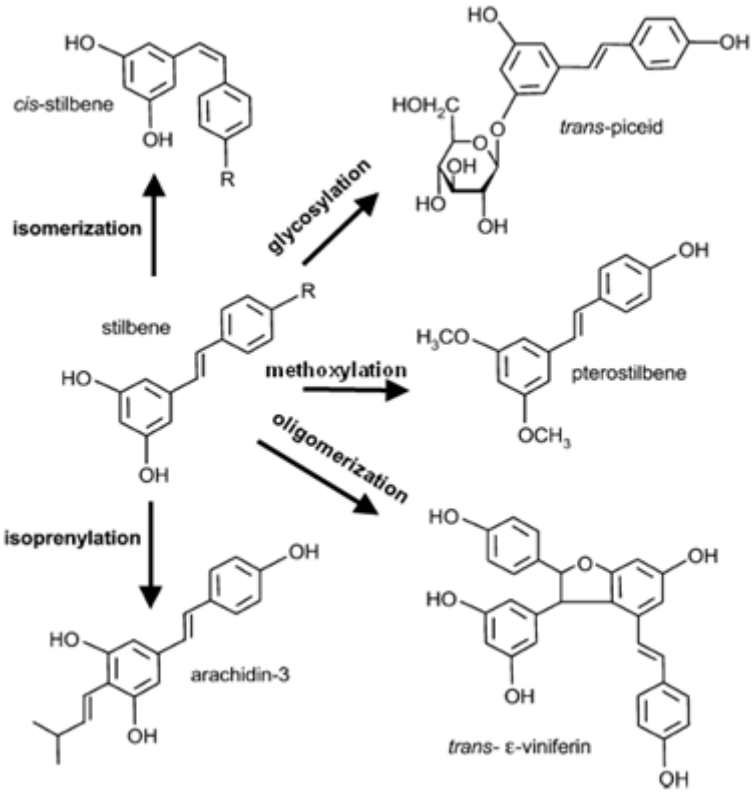


Fig. 14. Most common modifications of plant stilbenes (Chong et al, 2009).

3.4 Isocoumarin biosynthesis

Although considerable efforts have been devoted towards the chemical synthesis of isocoumarins, only a few reports about its biosynthesis have been published. Studies of isocoumarins biosynthesis were rather sporadic and there is still no well accepted pathway until now. Isocoumarins can derive from the acetate route as outlined in Fig. 15A. This sequence has been used to predict the correct structure of mullein before the exact structure was known (Birch et al., 1966). The biosynthesis of gentianine has also been considered using the carbohydrate approach (Fig. 15B). Thus, loss of water and carbon dioxide from starter to give the intermediate, followed by an oxidation-reduction sequence with loss of carbon dioxide could yield gentianine. An alternate biogenetic route is the stilbenecarboxylate route (Barry et al., 1964; Napolitano et al., 1997) (Fig. 15C). Some studies suggest that naturally occurring stilbenecarboxylate such as 3,4'-dihydroxystilbene-2-carboxylic acid and 3,3'-dihydroxy-4'-methoxysti1bene-2-carboxyIic acid may be the precursors of 3-aryl isocoumarins such as hydrangenol and phyllodulcinol. Stilbenes could arise from the condensation of a phenylacetic acid molecule such as 3,4-dihydroxyphenylacetic acid and an aromatic aldehyde such as 4,6-dihydroxyphthaldehydic acid, followed by decarboxylation and cyclization to isocoumarin (Seshad, 1957). Moreover, in view of the more recent developments in biogenetic sequences, this route has been mentioned many times. Even more interesting is that two new natural isocoumarins, cajanuslactone A and cajanuslactone C (Fig. 8), which have been isolated from the leaves of *C. cajan*, are

probably associated with the biosynthesis of stilbenecarboxylates in *C. cajan* and the biosynthesis of cajanuslactones are proposed to use a polyketide pathway (Fig. 16).

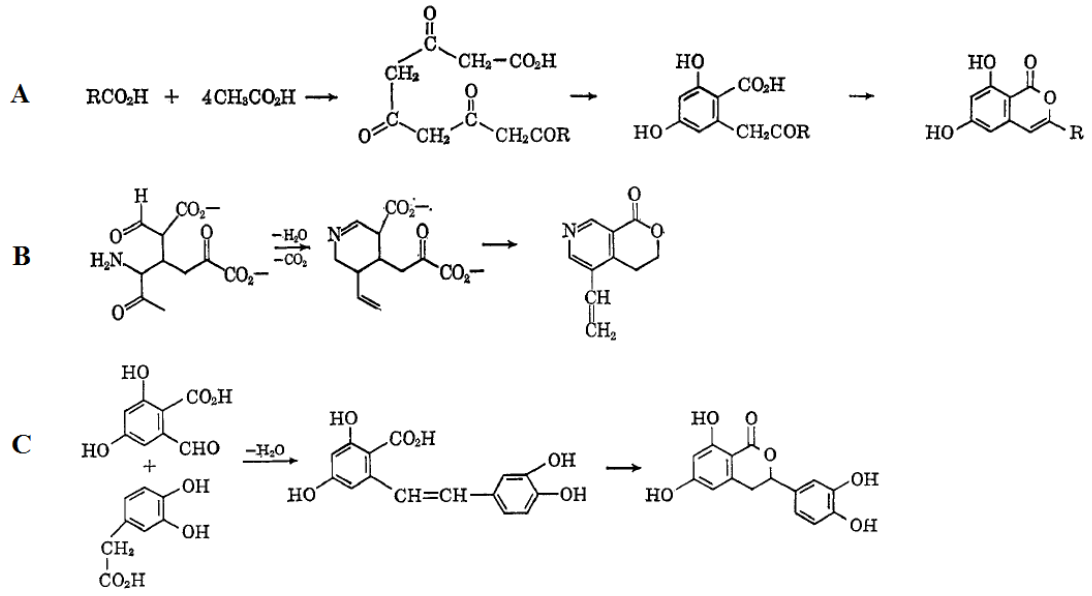
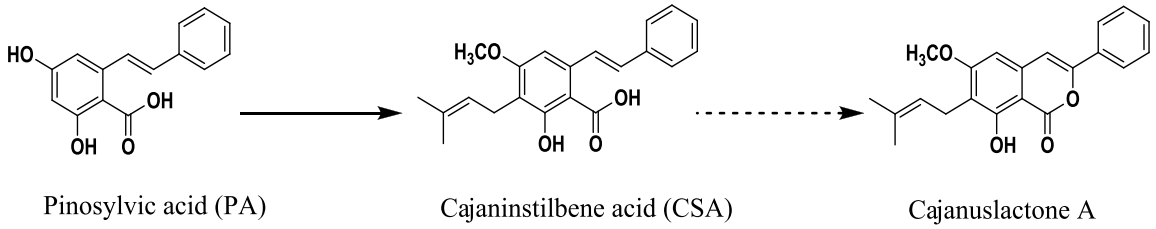


Fig. 15. Postulated biosynthetic pathways related to some isocoumarin derivatives
A: The acetate route, B: The carbohydrate route, C: The stilbenecarboxylate route.



Pinosylvic acid (PA) Cajaninstilbene acid (CSA) Cajanuslactone A

Fig. 16. Proposed cajanuslactone A biosynthetic pathway.

3.5 Stilbenecarboxylate biosynthesis

Although polyphenolics display an enormous chemical diversity, stilbenecarboxylates seem to constitute a rather restricted group of molecules. In earlier studies, stilbenecarboxylate once thought to be a derivative of stilbene. In recent years, labeling experiments and the incorporation of stilbenecarboxylates into well-characterized downstream natural products demonstrate that the isolation of in vivo stilbenecarboxylate is neither an artifact nor a by-product of stilbene biosynthesis (Shibuyaet al., 2002). These natural products presumably result via a nondecarboxylative STS-like (C2→C7 aldol) type III PKS reaction mechanism. The same reaction pathway, but proceeding from an aliphatic hexanoyl starter, is likely to be responsible for the carboxylate-bearing resorcinol moiety of tetrahydrocannabinol (THC) compounds in *Cannabis*. While candidate type III PKS enzymes have been cloned from some species, in vitro reconstitution of stilbenecarboxylate biosynthesis has been problematic at best.
Two independent attempts to clone and characterize stilbenecarboxylate synthases have

been published, with each group finding only one candidate (i.e., non-CHS) type III PKS enzyme (Akiyamaet al., 1999). Ebizuka's group, searching for hydrangic acid synthase of *Hydrangea macrophylla* var. thunbergii (Fig. 17A), cloned an enzyme that in vitro efficiently produced *p*-coumaroyltriacetic acid lactone (CTAL), a minor tetraketide CHS derailment product, from *p*-coumaroyl-CoA and three molecules of malonyl-CoA (Akiyamaet al., 1999). This enzyme was named *p*-coumaroyltriacetic acid synthase (CTAS) (Fig. 17A). Schröder's group independently observed the same lactone product specificity with the only apparent non-CHS type III PKS in *Hydrangea macrophylla* (garden hortensia), a species containing both hydrangic acid and the similar lunularic acid (Fig.17B). However, when primed with dihydro-*p*-coumaroyl-CoA as a starter, this enzyme produced nearly 50% 5-hydroxylunularic acid, a stilbenecarboxylic acid, in addition to triketide and tetraketide lactones. Notably, this group also detected 5-hydroxylunularic acid when a pine CHS was given the dihydro-*p*-coumaroyl-CoA starter, while a pine STS given the same starter failed to produce a stilbenecarboxylate. These results are extremely puzzling if one assumes stilbenes and stilbenecarboxylates are produced by a conserved biosynthetic mechanism.

Recently, another unique biosynthetic route to stilbenecarboxylate has been published as well. The Page group attempted to identify the PKS that synthesizes olivetolic acid (OA) which is an alkylresorcinolic acid that forms the polyketide nucleus of the cannabinoids (Gagne et al., 2012). They observed only olivetol and lactone products specificity with a STS type III PKS in *Cannabis sativa* (marijuana, hemp; Cannabaceae), a species containing Δ 9-tetrahydrocannabinol (THC) and other cannabinoids (Fig. 17C). However, they found a polyketide cyclase-like enzyme that could assist in OA cyclization. They showed that a type III PKS from cannabis trichomes requires the presence of a polyketide cyclase enzyme, olivetolic acid cyclase (OAC), which catalyzes a C2–C7 intramolecular aldol condensation with carboxylate retention to form OA. OAC is a dimeric $\alpha+\beta$ barrel (DABB) protein that is structurally similar to polyketide cyclases from *Streptomyces* species. This identification of OAC reveals a unique biosynthetic route to stilbenecarboxylate in which cyclases function cooperatively with type III PKSs to generate carbon scaffolds.

Based on a simple set of natural stilbenecarboxylates in *C.cajan* and that such stilbenecarboxylate synthase (STCS) activities have never been demonstrated in vitro, we became interested in these predicted polyketide synthase reactions. As outcome of this research, maybe the assignment to the family of CHS-related enzymes would add a new reaction type performed by these proteins and provide, as a consequence, possibilities to synthesize new products.

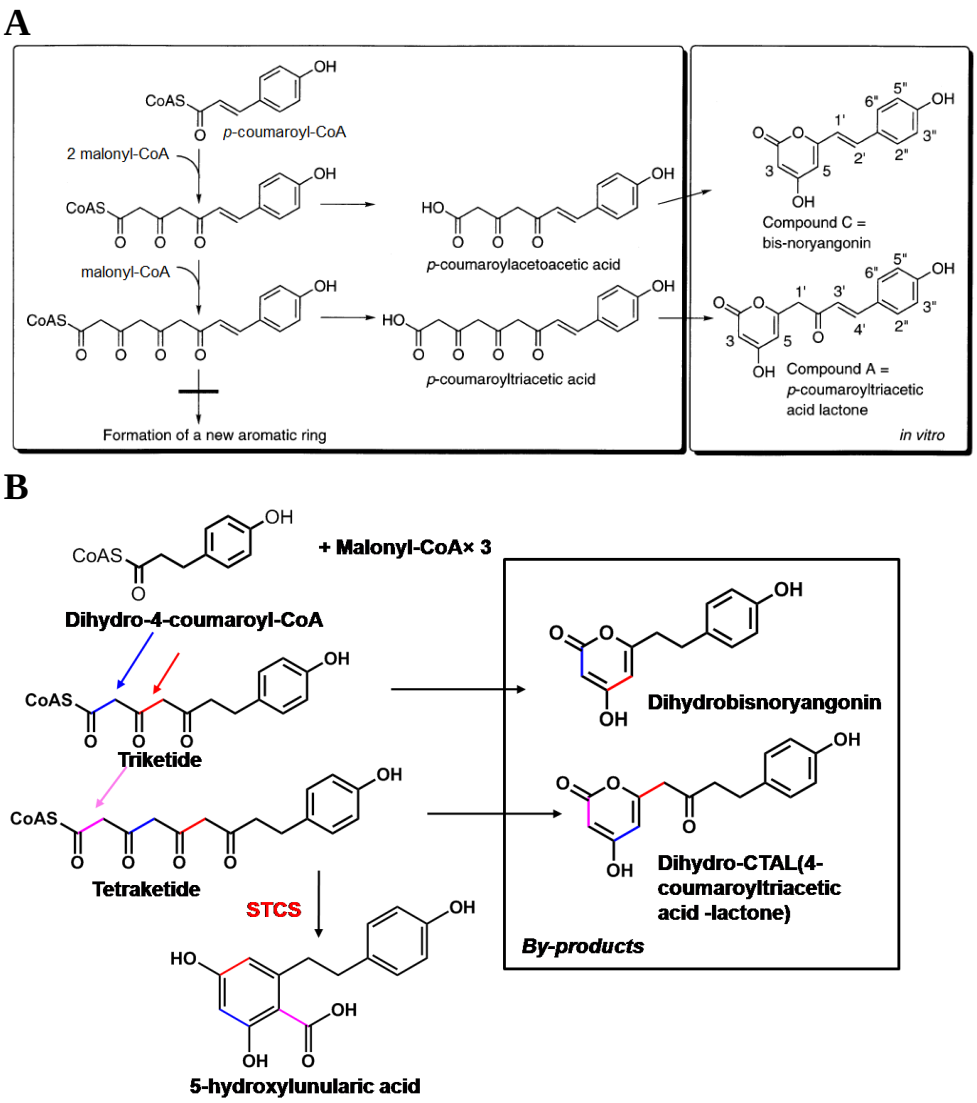


Fig. 17. Established biosynthetic pathways of stilbenecarboxylates. A, B: 5-hydroxylunularic acid formation; C, olivetolic acid formation

3.6 Postulated biosynthetic pathways of main active constituents in *C. cajan*

In contrast to the extensive phytochemical studies there is no any data on the biosynthesis of secondary metabolites in *C. cajan*. In the four above-mentioned secondary metabolite groups (flavonoids, stilbenoids, isocoumarins and stilbenecarboxylates), the key steps of their biosyntheses are catalysed by type III polyketide synthases. The two well-studied type III PKSs, chalcone (CHS) and stilbene (STS) synthases are the key enzymes in the biosynthesis of flavonoids and stilbenoids, respectively. They have been extensively studied in several plants. Precursor feeding studies and mechanistic rationalization suggest that stilbenecarboxylates might also be synthesized by plant type III polyketide synthases, which are stilbenecarboxylate synthases (STCS). However, the detection of STCS activity and its molecular characterization has not ended the debate about the origin of the terminal carboxyl group of stilbenecarboxylate. *C. cajan.* contains a good amount of stilbenecarboxylate such as cajaninstilbene acid, isocajaninstilbene acid and 4-*O*-methylpinosylvic acid. More interestingly, a series of products which belong to stilbenes and isocoumarins may associate with the formation of stilbenecarboxylates. They are likely to have derived skeletons with similar substitution patterns. In addition, their simultaneous production in the plant was detected in extensive studies. According to all these findings, we became interested in these predicted polyketide synthase reactions. The postulated biosynthetic pathway is discussed in Fig.18.

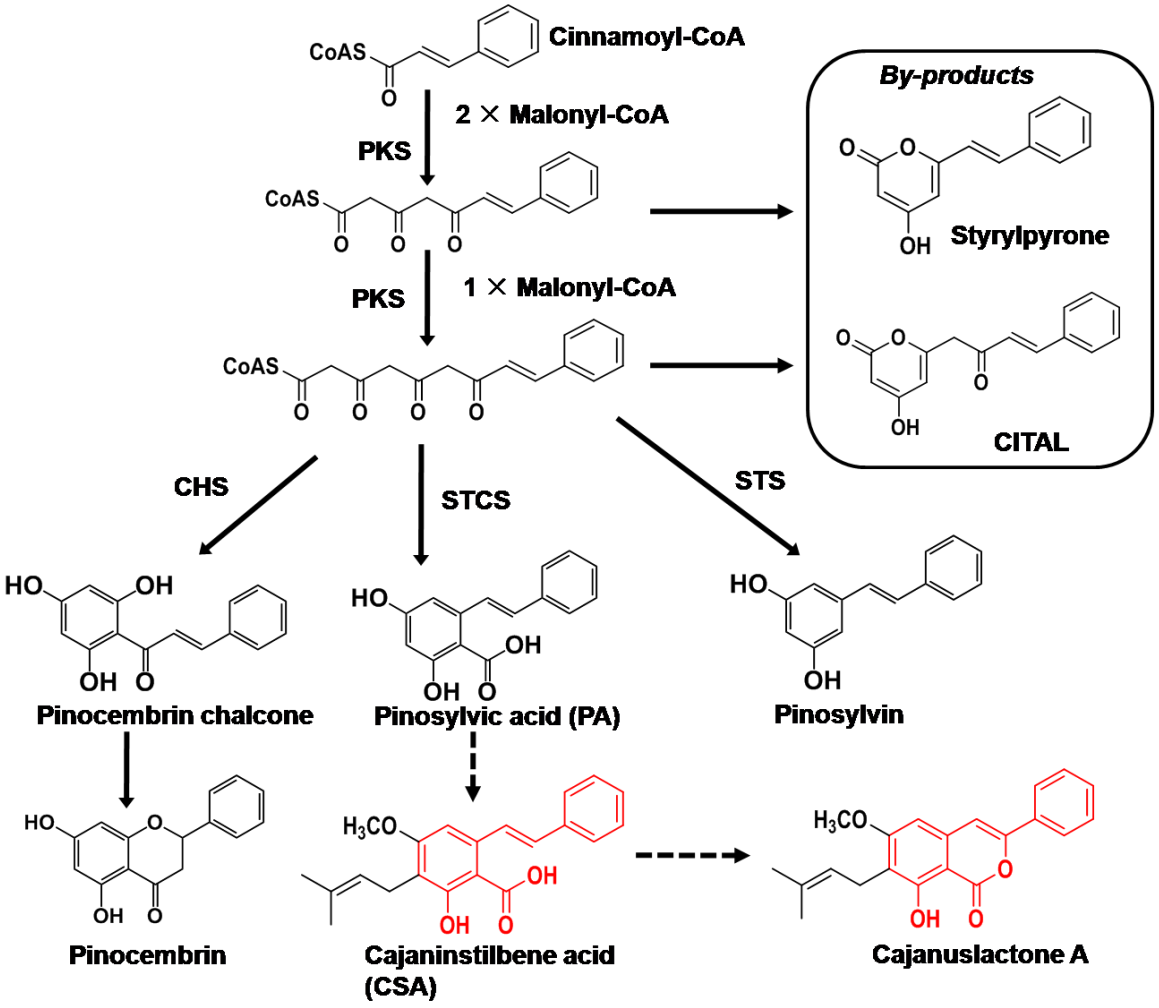


Fig. 18. Proposed biosynthetic pathwayof stilbenecarboxylate and related secondary metabolites in *C. cajan*.

Precursor feeding studies indicated that the backbone of stilbenecarboxylate is synthesized via the shikimate/malonate route like chalcones and stilbenes (Billek and Kindl, 1962; Ibrahim and Towers, 1962; Pryce, 1971). The enzyme reaction should involve a STS-type ring folding (C2→C7 aldol) without removal of the terminal carboxyl group. The formation of the natural products also requires a reducing step, but it is unknown at which level it occurs. The biosynthesis of cajanuslactone A is proposed to use a polyketide pathway (Fig. 18). The first intermediate is pinosylvic acid (PA), which has been postulated to be synthesized by a new type III polyketide synthase (PKS), stilbenecarboxylate synthase (STCS), which uses cinnamoyl-CoA as starter. The postulated CHS and STS reactions are also included in this pathway.

4. Research strategies and objectives

C. cajan not only contains stilbenecarboxylate but also other kinds of secondary metabolites with excellent biological activities such as flavonoids, stilbenes and isocoumarins. Although these constituents were extensively studied on the phytochemical level, their biosynthesis was poorly investigated. Especially for STCS activities, the enzymatic basis for this reaction remains unclear. However, the great amount of stilbenecarboxylates and their derivatives (e.g. cajaninstilbene acid) in *C. cajan*, which are probably derived from cinnamoyl-CoA (Fig. 10), provide us with a new resource to do further studies on the STCS reaction. STCS from *C. cajan* is likely to be a type III PKS that is the key enzyme of the biosynthetic pathway of stilbenecarboxylate and presumably also isocumarins. The aim of this work is to study the biosynthesis of the main secondary metabolites in *C. cajan* at both the enzyme and the gene level, especially for the related type III PKSs which are involved in these metabolic processes. A facile plant material system with high enzyme expression for biochemical and molecular studies should be established. Moreover, enzyme assays by incubating possible substrates and intermediates with different protein preparations (crude protein extracts) aimed to elucidate the detailed biosynthetic steps and the enzymes involved. Furthermore, isolation, cloning, and functional characterization of cDNAs encoding thedetected metabolizing enzymes are essential prerequisites for future manipulation of the expression of these genes. In addition, the assignment to the family of CHS-related enzymes might add a new reaction type performed by these proteins, thereby providing opportunities to synthesize new products.

II. Materials and Methods

1. Materials

1.1 Plant material

Studies were performed with different tissue cultures of *C.cajan*. Seeds were obtained from the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT). The seeds of *C.cajan* were sterilized as described in (II.2.1.1). *In vitro* plants were grown on solid Murashige and Skoog (MS) medium containing 3% sucrose as carbon source without any phytohormones under controlledconditions at 25°C and 16h light/ 8h dark cycle. The medium was refreshed ateight-week-intervals. Leaves of in vitro plants were used as anexplant for callus induction. Callus was grown on solid MS medium supplemented with hormons (see II.2.1.2 for details). The tissue cultures were maintained by regular subculturing at intervals of 3 weeks. Callus was suspended inliquid MS medium supplemented with phytohormons (II.2.1.3). The resulting cell suspension cultures were propagated in 300 ml Erlenmeyer flasks, which were maintained on an incubator shaker at 100 rpm and 25°C in the light. For following studies, 3 g cells were inoculated into 50 ml liquid MS medium at 21-day-intervals.

1.2 Chemicals

The suppliers of chemicals used in the work are given as follows. Deionized water supplied by a Milli-Q water purification system (Sartorius, Germany) was used in preparing all aqueous solutions used in the study. All solutions were autoclaved by 121°C for 20 min. Solutions of thermolabile compounds were filter-sterilized and added to autoclaved solutions under sterile conditions.

Chemical	Supplier
Phytohormones	
2,4-Dichlorophenoxyacetic acid (2,4-D)	Fluka
1-Naphthylacetic acid (NAA)	Fluka
Indole-3-acetic acid (IAA)	Fluka
Indole-3-butyric acid (IBA)	Acros
6-Benzylaminopurine (BAP)	Fluka
Kinetin (KN)	Applichem
Elicitors	
Methyl jasmonate	Serva
Chitosan	Roth
Yeast	Applichem
Sodium chloride	Roth
Antibiotics	
Ampicillin	Roth

Chloramphenicol	Fluka

Reagents used for crude protein extraction

Polyclar®AT	Serva
Seasand	Roth
Hydroxyethylpiperazin ethan sulfonic acid(HEPES)	Sigma-Aldrich
Sodium citrate	Roth
Sodium acetate	Roth
Sodium phosphate	Roth
Sodium borate	Roth
Tris–HCl	Roth
Dithiothreitol (DTT)	Applichem

Stationary phases used for protein affinity purification, desalting andconcentration

Nickel-nitrilotriacetic acid (Ni-NTA) agarose	Qiagen
PD10-cartridge Sepharose G-25 columns	GE Healthcare
Vivaspin 6 centrifugal concentrators	Biotech

Chemicals for enzyme assays

Malonyl-CoA	Sigma-Aldrich
Cinnamoyl-CoA	Sigma-Aldrich
p-Coumaroyl-CoA.	Sigma-Aldrich
Dihydro-*p*-coumaroyl-CoA	Synthesized
Dihydro-cinnamoyl -CoA	Synthesized
Acetyl-CoA	Sigma-Aldrich
Benzoyl-CoA	Sigma-Aldrich
m-Cinnamoyl -CoA	Self-synthesized
o- Cinnamoyl -CoA	Self-synthesized
Isobutyryl-CoA	Sigma-Aldrich
Coenzyme A	Sigma-Aldrich
Pinosylvin	Sigma-Aldrich
Resveratrol	Sigma-Aldrich
Bisnoryangonin	Sigma-Aldrich
Pinocembrin	Sigma-Aldrich
Naringenin	Sigma-Aldrich

Reagents used to stop enzymatic reaction

Trichloroacetic acid	Fluka
Glacial acetic acid	Roth

Reagents for biochemistry and molecular biology

DMSO	Fluka
IPTG	Sigma-Aldrich
X-Gal	Sigma-Aldrich
dNTPs	Fermentas
Glutathione (Reduced form)	Sigma
TRIS-HCl	Roth

Imidazole	Roth
Reagents for GC-MS derivatization	
N-methyl-N-(tri-methylsilyl)-trifluoroacetamide (MSTFA)	ABCR
Methoxyamine hydrochloride	Sigma-Aldrich
Reagents for gel electrophoresis	
peqGOLD Universal Agarose	Peqlab
Ethidium bromide	Roth
Urea	Bio-Rad
TEMED	Bio-Rad
Acrylamide / Bisacrylamide 30%	Bio-Rad
Ammonium persulfate	Roth
EDTA	Sigma-Aldrich
Formamide	Sigma-Aldrich
SDS	Roth
ß-mercaptoethanol	Fluka
Bromophenol blue	Sigma
Coomassie-blue R 250 and G 250	Merck
Solvents for HPLC	
Methanol	Fisher-Scientific
Orthophosphoric acid	Roth
Trifluoroacetic acid	Roth
Formic acid	Roth
Acetonitrile	Fisher-Scientific
Ladders	
PageRuler Unstained Protein Ladder (10-170 kDa)	Fermentas
Gene Ruler DNA Ladder Mix	Fermentas

1.3 Nutrient Media

1.3.1 Nutrient medium for plant tissue culture

MS medium (Murashige and Skoog, 1962)

	Stock solution ingredients		Supplier	For 1 litermedium
Macro elements	NH_4NO_3	33 g	Roth	100 ml
	KNO_3	38 g		
	$CaCl_2$. $2H_2O$	8.8 g		
	$MgSO_4$. $7H_2O$	7.4 g		
	KH_2PO_4	3.4 g		
	H_2O to	1 liter		
Micro elements	H_3BO_3	6.2 g	Roth	1 ml
	$MnSO_4$. H_2O	16.9 g		
	$ZnSO_4.7H_2O$	8.6 g		
	KI	0.83 g		

	Na_2MoO_4. $2H_2O$ $CuSO_4$. $5H_2O$ $CoCl_2$.$6H_2O$ H_2O ad	0.25 g 0.025 g 0.025 g 1liter		
Ferric salt	Na_2EDTA .$2H_2O$ $FeSO_4$. $7H_2O$ H_2O ad	41.3 g 27.8 g 1 liter	Roth Merck	5 ml
Vitamins	Nicotinic acid Pyridoxin-HCl Thiamine HCl Glycin Myo-Inositol H_2O ad	0.05 g 0.05 g 0.01 g 0.2 g 10 g 1 liter	Merck Merck Serva Roth Sigma	10 ml
Hormones	NAA/ Ethanol BAP/ Ethanol	1 mg/ml 1 mg/ml	Fluka Fluka	100 µl 300 µl
Sucrose			Fluka	30 g
pH				5.7
Solid LS(for callus cultures)			Roth	0.8% Agar

1.3.2 Bacterial culture media and reagents

Medium	Components	
LB medium For solid medium	Bacto-pepton Yeast extract NaCl Agar	10 g/l 5 g/l 10 g/l 1.5%
SOC-Medium	Peptone from Casein Yeast extract 1 M NaCl 1 M KCl Autoclave, then add: Sterile filtered solution of 2 M Mg^{2+} 2 M glucose	20 g/l 5 g/l 10 ml/l 2.5 ml/l 10 ml/l 10 ml/l
For preservation of bacterial culture	Bacterial culture in LB medium Autoclaved solution of : Glycerin (Roth):LB (2:8)	250 µl 750 µl
Antibiotic	Ampicillin (100mg/mL) Chloramphenicol (30mg/mL)	100 µg/ml medium 30 µg/ml medium
For induction of protein expression	IPTG (0.5M) (filter-sterilized and freshly prepared)	1 mM Final concentration
For blue/white screening *X*-Gal	5-Bromo-4-chloro-3-indolyl-*β*-Dgalactopyranosid (*X*-Gal 40mg) +N,N`-Dimethylformamid (1 ml)	40 µl per petridish

1.4 Buffers and solutions

1.4.1 Buffers and solutions for gel electrophoresis

DNA- agarose electrophoresis

50X TAE buffer	Tris-HCl EDTA Adjust pH to 8 with glacial acetic acid	2 M 0.05 M
Agarose-gels were prepared at concentrations of 0.6-2% agarose in electrophoresis buffer (1x TAE)		

Protein electrophoresis (SDS-PAGE)		
Stacking gel (5%)	dH_2O	3.4 ml
	0.5 M Tris-HCl (pH 6.8)	0.63 ml
	30% Acrylamide/Bis.	0.83 ml
	10% (w/v) SDS	0.05 ml
	10% (w/v) APS	0.05 ml
	TEMED	5 µl
Separating gel (12%)	dH_2O	3.3 ml
	1.5 M Tris-HCl (pH 8.8)	2.5 ml
	30% Acrylamide/Bis.	4 ml
	10% (w/v) SDS	0.1 ml
	10% (w/v) APS	0.1 ml
	TEMED	4 µl
TEMED is used with APS to catalyze the polymerization of acrylamidewhen making polyacrylamide gels, which are used in gel electrophoresis for theseparation of proteins or nucleic acids.		
Protein loading Buffer (2 X)	dH_2O	2.7 ml
	0.5 M Tris-HCl (pH 6.8)	1ml
	Glycerin	2 ml
	10% (w/v) SDS	3.3 ml
	β-mercaptoethanol	0.5 ml
	Bromophenol blue 0.5% (w/v)	0.5 ml
10x Electrophoreticbuffer	Tris-HCl	15 g
	Glycin	72 g
	Na-SDS	5 g
	dH_2O	ad 500 ml
Staining solution	Coomassie-blue R 250	125 ml
	Methanol	500 ml
	Acetic acid	100 ml
	dH_2O	ad 1000 ml
Coomassie-blue stock solution	Coomassie-blue R 250	5 g
	dH_2O	ad 500 ml
Destainingsolution	Methanol	200 ml
	Acetic acid	76 ml
	dH_2O	1000

1.4.2 Buffer used for crude protein extraction

HEPES buffer	100 mM	**pH 7**
Sodium citrate buffer	100 mM	**pH 4**
Sodium acetate buffer	100 mM	**pH 5**
Sodium phosphate buffer	100 mM	**pH 7**
Tris–HCl buffer	100 mM	**pH 8**
Sodium borate buffer	100 mM	**pH 9**

1.4.3 Buffers for affinity purification of fusion protein

His_6-tagged fusion protein

Lysis buffer pH 8	50 mM Na_2HPO_4 30 mM NaCl 20 mM Imidazole Water	3.44 g 0.87 g 0.68 g ad 500 ml
Washing buffer pH 8	50 mM Na_2HPO_4 1.5 mM NaCl 50 mM Imidazole Water	3.44 g 0.04 g 1.7 g ad 500 ml
Elution buffer pH 8	50 mM Na_2HPO_4 300 mM NaCl 250 mM Imidazole Water	3.4 g 8.76 g 8.51 g ad 500 ml

1.4.4 Buffers used for enzyme assays

KH_2PO_4 buffer 100 mM, the pH was adjusted with concentrated NaOH, autoclaved andstored at room temperature.

Tris-HCl buffer 100 mM, the pH was adjusted with concentrated NaOH, autoclaved and stored at room temperature.

1.4.5 Solution for protein estimation

Bradford-dye solution	Coomassie-brilliant blue G-250 Ethanol 96% *o*-phosphoric acid 85% w/v Water	100 mg 50 ml 100 ml ad 1000 ml
Dissolve well Coomassie®-Brilliant G250 in ethanol, add orthophosphoric acid and make volume up to 1 L with water. Filter the solution through filter paper (Whatman No. 1) until no blue colour can be seen. Keep at 4°C in amber glass bottle.		

1.4.6 Solutions for PD10 washing and Ni-NTA agarose regeneration

PD10 washing solution NaOH (0.15 M)	Five column volumes of NaOH cleaning solution followed by five volumes of distilled water until getting a neutral eluent.

Ni-NTA agarose washing and regeneration solution	0.2 M acetic acid, 30% glycerol, deionized water Wash in the same order

1.4.7 Buffers used for plasmid isolation (mini prep)

Buffer 1	Tris-HCl	50 mM, 1.5 g/250 ml
	EDTA	10 mM, 0.93 g/250 ml
	RNase A	100 µg/ml
	Adjust to pH 8 with HCl	
	RNase A was freshly added prior to use	
Buffer 2	NaOH	0.2 M, 2 g/250 ml
	SDS	1% (w/v), 2.5 g/250 ml
Buffer 3	Potassium acetate	2.55, 62.57 g/250 ml
	Adjust to pH 5.5 with glacial acetic acid	

1.5 Materials used for molecular biology

1.5.1 Host cells and cloning vectors

1.5.1.1 Host cells (competent *E. coli*)

E.coli	**Purpose**	**Genotype**
DH5α	Deliver high yield plasmid preparation for downstream applications, chemically competent	*F' φ80δlacZ9M15 end A1hsdR17(rk-mk+)supE44thi-1 λ -gyrA96 relA19(lacZYA-argFV169) deoR*
BL21(DE3)pLysS	For protein overexpression (Invitrogen), chemically comptent	*F⁻ompT hsdSB (rB⁻mB⁻) gal dcm (DE3) pLysS (CamR)*

1.5.1.2 Vectors

pRSET-B vector (protein expression)	2,9 kb expression vector with N-terminal His_6-tag and ampicillin resistance gene	Invitrogen

1.5.2 Primers

All primers were synthesized in HPSF (High Purity Salt Free) quality at MWG-Biotech AG (Ebersberg, Germany)

Label	**Sequence**
STCS gene-specific primers	
PAS1for *Nhe* I	5`-AGTGCTAGCATGGCATATTTGCAAGAAATACGTGAG-3`
PAS 1 rev *Xho* I	5`-AGTCTCGAGCTAATGGCTTGGGTCGATGGTG-3`

PAS2 for *Nhe* I	5`-AGTGCTAGCATGGCATATTTGGAGGAAATACGTGAG-3`
PAS 2 rev *Xho* I	5`-AGTCTCGAGCTAATAATGGCTTGGATCGATGGTG-3`
PAS3 for *Nhe* I	5`-AGTGCTAGCATGGCATACTTGCAGGAAATACATGAG-3`
PAS 3 rev *Xho* I	5`-AGTCTCGAGCTAGCTATAGTGGCCTGAGTCTAAG-3`
CHS gene-specific primers	
C.CHS 1 for *Nhe* I	5`-AGTGCTAGCATGGTGTCTGTTGATGAGATCCGCAAG-3`
C.CHS 1 rev *Eco* RI	5`-AGTGAATTCTCAAGCTGCAACACTATGGAGAACAACG-3`
C.CHS 2 for *Nhe* I	5`-AGTGCTAGCATGGCAACGGTCGATGAGATTCACAG-3`
C.CHS 2 rev *Eco* RI	5`-AGTGAATTCTTAGAGTGCTACACTATGAAGAACAATAGTCTC-3`
C.CHS 3 for *Nhe* I	5`-AGTGCTAGCATGGTGAGTGTTGAAGATATCCGAAAGGC-3`
C.CHS 3 rev *Eco* RI	5`-AGTGAATTCTTAGACAGTGACACTGCGGAGCACAA-3`
C.CHS 4 for *Nhe* I	5`-AGTGCTAGCATGGCCAGCGTTGCTGAGATCC-3`
C.CHS 4 rev *Eco* RI	5`-AGTGAATTCTTATATGGCCACACTGTGCAGAACAACAG-3`
C.CHS 5 for *Nhe* I	5`-AGTGCTAGCATGGTGACGGTGGAGGAAATCCG-3`
C.CHS 5 rev *Eco* RI	5`-AGTGAATTCTCAACCCTCCAAGGGAACGCTATGAA-3`
SAC gene-specific primers	
DAC 1 for *Nhe* I	5`-AGTGCTAGCATGGGGGAATTCAAACACTTTGTAATTG-3`
DAC 1 rev *Kpn* I	5`- AGTGGTACCTCATGCTGGGGCCTTGACAAGA-3`
DAC 2 for *Nhe* I	5`-AGTGCTAGCATGGGGACTTTCAATCACTATGTGG-3`
DAC 2 rev *Kpn* I	5`-AGTGGTACCTCATGCTGGTGCTTTGACAAGCTTA-3`
DAC 3 for *Nhe* I	5`-AGTGCTAGCATGGAAGAAGCAAAGGGATTAGTGAAC-3`
DAC 3 rev *Kpn* I	5`-AGTGGTACCYTACAATTTAACAAYGGTAGGCTTATAG-3`
DAC 4 for *Nhe* I	5`-AGTGCTAGCATGGAGGAAGCAAAGGGATTAGTGAAA-3`
DAC 4 rev *Kpn* I	5`-AGTGGTACCTCAAAGGTTTACAGTAGTAGGCTTG-3`
DAC 5 for *Nhe* I	5`-AGTGCTAGCATGGAGGAAGCAAAAGGATTGGTGAAA-3`
DAC 5 rev *Kpn* I	5`-AGTGGTACCYTACAATTTAACAAYGGTAGGCTTATAG-3`
BEC 1 for *Nhe* I	5`-AGTGCTAGCATGGAGGCCCTCACTTTCACTG-3`
BEC 1 rev *Kpn* I	5`-AGTGGTACCTCAAGCATAGGCCTCTGGGTTTG-3`
BEC 2 for *Nhe* I	5`-AGTGCTAGCATGGGTGTTTTCASATTCGAGGATG-3`
BEC 2 rev *Kpn* I	5`-AGTGGTACCCTAGTAATCAGGATTGGCCAAAAGG-3`
BEC 3 for *Nhe* I	5`-AGTGCTAGCATGGGTGTTTTCTCATTTGAGGATGAAAC-3`
BEC 3 rev *Kpn* I	5`-AGTGGTACCTCAGTAATSAGGATTGGCCAAAAGG-3`
BEC 4 for *Nhe* I	5`-AGTGCTAGCATGGGTGTTTTCACATTCGAGGATG-3`
BEC 4 rev *Kpn* I	5`-AGTGGTACCTCAGTAATCAGGATTGGCCAAAAGG-3`
BEC 5 for *Nhe* I	5`-AGTGCTAGCATGGGTGTTTTCACTTGTGAAGTCG-3`
BEC 5 rev *Kpn* I	5`-AGTGGTACCTTAGTTGTAATCAGGATTAGCCAGAAGG-3`
CHC for *Nhe* I	5`-AGTGCTAGCATGGCATTCATGCGCTTCCCTTTC-3`
CHC rev *Kpn* I	5`-AGTGGTACCCTAGAGGTATGTAGCTACGTTGTTC-3`
Vector-specific primers (for sequencing the DNA insert)	
pRSET B	for: 5`-GAG ACC ACA ACG GTT TCC CTC-3`
	rev: 5`-CTA GTT ATT GCT CAG CGG TGG-3`

for: forward; rev: reverse; restriction sites are underlined. PAS: Pinosylvic acid synthase; C.CHS: *C. cajan* CHS; SAC: stilbenecarboxylate cyclase; DAC: DABB type cyclase; BEC: Betv1-like type cyclase; CHC: CHI-like type cyclase

1.5.3 Enzymes

Name	Purpose	Supplier
RevertAid H Minus Reverse transcriptase	Reverse transcription	Thermo Scientific
Phusion Hot Start II High fidelity DNA polymerase	High fidelity amplification of DNA	Thermo Scientific
Taq-DNA Polymerase (Dream *Taq*)	DNA polymerases	Fermentas
*Eco*R1, *Nhe*I, *Kpn*I, *Xho*I	Restriction enzymes	Thermo Scientific
T4-DNA ligase	Ligation	Thermo Scientific
RNase A	Digestion of RNA by plasmid isolation	Thermo Scientific
RNase-free DNase I	Digestion of genomic and contaminating DNA by RNA isolation	Qiagen

1.5.4 Kits

RNA isolation	Rneasy®Plant mini kit	Qiagen
DNA isolation	Dneasy™plant maxi kit	Qiagen
Purification of DNA from PCR, restriction product, orfrom gel	innuPREP DOUBLE pure kit	Analytic Jena biosolutions

1.6 Equipments

Equipment	**Model**	**Company**
Water purification system	Arium 611 VF	Sartorius, Germany
Autoclave	Vx-120	Systec GmbH Laborsystemtechnik
Balance	Small and large scale	Satorius,Germany
pH meter	Digital pH meter 325	WTW
Centrifuge	Universal 32R	Hettich
	Biofuge 13	Heraeus Sepatech
	Sigma 1-15K	Sigma
GC-MS	Gas chromatograph 6890	Agilent
	ZB5-MS column (30 m, 0.25 mm i.d, 0.25 μm ft)	Phenomenex, Aschaffenburg
	Mass spectrometer	JEOL
LC-MS (ESI-MS)	3200 Q TRAP	Appl. Biosystems
Scintillation System	LS 6500	Beckman Coulter
Spectrophotometer	Ultrospect 1000	Pharmacia Biotech
Vacuum concentrator	RVC 2-18	CHRIST
Incubator shaker	HT	Infors

Thermo block	Dri-Block DB-3D	Techne
HPLC	Elite LaChrom series L-2200 autosampler quaternary L-2130 Pump equipped with low-pressure gradient L-2455 diode array detector EZChrome Elite software HyperClone ODS column (C18, 150 x 4.6 mm, 5µm)	VWR-Hitachi Phenomenex
	Agilent 1200 series Agilent HP 1200 binary pump Agilent HP 1200 autosampler Agilent 1200 variable wavelength detector. Gina Star 4.06 software. HyperClone ODS column (C18, 150 x 4.6 mm, 5µm)	Agilent
	Ramona Star radiodetector with soild scintillation analysis cell quartz tube (5.5 mm, 0.37 ml, particle size 45-63 µm)	Raytest GmbH
	Lichrospher 100 RP 18 E (150 x 4.6 mm, 5 µm)	WICOM
Balance	LA 230S	Satorius
Gel documentation	MultiImage^{TM}Light Cabient	Alpha In. Corp
Heating Circulator water bath	MW-4	Julabo
PCR cycler	T-Proffessional Gradient	Biometra
Clean bench	LaminarAir HLB 2472 Laminar Air HBB 2460	Heraeus Heraeus
Magnetic rotator	VF2	IKA-Labortechnik (Janke & Kunkel)
Electrophoresis	Mini-Sub Cell Sub Cell-GT Protein Chamber	BioRad BioRad Biometra
Power supply	Standard Power Pack P25 Power Pack 300	Biometra BioRad
Ultrasonic-Cell-Disruptor	Sonifier 250	Branson (G.Heinemann)
-80°C freezer	Hera Freeze	Heraeus

2. Methods

2.1 Establishment and selection of plant tissue cultures

2.1.1 Seed sterilization and germination

Mature seeds of *C. cajan* were obtained from the gene bank of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). Seeds were surface sterilized in 70% EtOH for 2 min and then in 6% v/v sodium hypochlorite solution for 5 min. They were finally rinsed with sterile distilled water several times till all traces of sodium hypochlorite were eliminated. Subsequently, the seeds were soaked in sterile distilled water for 18 h at 25±2°C in the dark, drained aseptically, and washed twice in sterile double-distilled water. Pre-soaked seeds were germinated aseptically in Petri dishes containing 25 ml of hormone-freeMS (Murashige and Skoog, 1962) medium supplemented with 3% (w/v) sucrose and solidified with 0.8% agar (w/v). The pH of MS medium was adjusted to 5.7 before sterilization by autoclaving at 121°C for 20 min. The pH of MS medium was confirmed after sterilization and kept in a growth chamber at 25 ± 2°C under a 16 h photoperiod till the seeds germinated to produce seedlings. Seven to 10-day old *in vitro*-grown seedlings, which had produced the first two fully expanded simple leaves and another well-developed trifoliate leaf besides the shoot tip, were used as the source of leaf explants for initiating callus.

2.1.2 Callus induction

Leaf explants were excised from 7-day-old seedlings and cut into 0.5 cm^2 segments which includedthe midrib. The explants were transferred into culture jar with MS medium containing 3% sucrose and 0.8% agar, supplemented with different hormone (auxins and cytokinins) concentrations for callus induction. The auxins tested were 2,4-dichlorophenoxyacetic acid (2,4-D) and α-naphthaleneacetic acid (NAA). The cytokininstested were kinetin (KT) and 6-benzylaminopurine (BAP) at the different concentrations. Batch cultures were incubated either inthe culture room at 25 ± 2 °C under a 16 h photoperiod or in the dark. The pH was adjusted to 5.7 prior to autoclaving and all experiments were repeated three times.

2.1.3 Establishment of cell suspension cultures

Ten-day-old greenish-white friable calli (3 g fresh weight) from MS medium were aseptically transferred to 250-ml Erlenmeyer flasks containing 50 ml MS liquid medium with 1 mg/l of 2,4-D and agitated on a gyratory shaker (100-110 rpm). Cultures were maintained at 25 ± 2 °C and 16-h-day photoperiod. Cell suspensions were subcultured by adding 20 ml of the old suspension to 50 ml of fresh medium at two-week-intervals. Cultures were filtered through 125 μM stainless steel sieves to separate individual cellsand small cell clumps. Various concentrations of the cytokinins were tested to promote the growth of cell cultures. After 3 subcultures, the dry weight of the cultured cells was measured and they were used in the following tests. All experiments were repeated three times.

2.1.4 Extraction of compounds

Extractions of compounds from *in vitro* leaves, callus and cell suspension cultures of *C. cajan* were carried out as described by Liu et al. (2007) with slight modifications. To 3 g of lyophilized and powdered plant material, 4 ml EtOH:H_2O (8:2, v/v) was added, vortexed

for 30 s and sonicated for 10 min. The mixtures were centrifuged at 5000 rpm, 4°C for 20 min. The EtOH:H_2O fractions were separated and evaporated. The extraction was performed twice. The extracts were resuspended in 1 ml of MeOH for the subsequent stilbenecarboxylates analyses.

2.1.5 Preliminary HPLC analysis of crude extracts

The crude extracts of the plant tissues were analyzed by an Agilent HPLC system (Agilent 1200 series) coupled with a Diode Array Detector (DAD) and equipped with Hyper Clone ODS column (C18, 150 x 4.6 mm, 5 µm). Themobile phase consisted of water containing 1% *ortho*-phosphoric acid, pH 2.5 (A) and methanol (B) with step gradient elution as follows:

Time [min]	**Methanol [%]**
0	65
3	65
10	85
25	95
27	100
29	100
30	65
35	65

The flow rate was 0.5 ml/min, the sample injection volume was 20 µl, and detection was at 259 nm.

2.1.6 Accumulation of target compounds in different organs of *C. cajan* after germination

In vitro plants were grown on solid Murashige and Skoog (MS) medium containing 3% sucrose as carbon source without any phytohormones under controlled condition at 25°C and 16h light/ 8h dark cycle. The medium was refreshed at eight-week-intervals. This experiment was done with *in vitro* plants, which were well grown in the same period. A total of 10 replicate cultures were used per treatment and the experiment was repeated twice. Different organs (leaves, stems and roots) of each repetition were taken under sterile conditions at different time points. Preparation of samples for HPLC analysis and quantification were carried out as described above. Dry weight of plant tissues was determined and the identity of the target compound was confirmed by UV spectroscopy, LC-MS and cochromatography with authentic reference compounds. Data represent three technical repeats for one biological sample.

2.1.7 Growth curves of suspension cultures

Growth curves were determined by inoculating 3 g fresh weight of *C. cajan* cell suspension cultures into 50 ml of fresh MS medium supplemented with 0.3 mg/l BA and 0.1 mg/l NAA in 250 ml flasks. The harvesting was performed in triplicate at intervals of 2 days. The cells were separated from the medium by suction filtration, weighted for fresh weight determination. Dry weight was determined after drying the cells at 50 °C in a hot air oven.

Cells were powdered, then extracted and subjected to quantitative analysis of target compounds.

2.1.8 Elicitation of cell suspension cultures

2.1.8.1 Elicitors preparation

Different elicitors were used for treatment of cell suspension cultures.
Preparation and storage were as follows:

Yeast extract	Stock solution of 300 mg/ml was prepared in water, filter sterilized by 0.20 μM filter and kept at 4°C.
Chitosan	Chitosan (a polymer of *β*-1,4-glucosamine residues) stock solution was prepared according to Pitta-Alvarez and Giulietti (1999). A concentrated chitosan solution was prepared by dissolving 125 mg chitosan in 1% (v/v) acetic acid, the final volume was made up to 100 ml. The pH was adjusted to 5.5 with 1N NaOH, sterilized by autoclaving at 121 °C for 20 min, and kept at 4°C.
Methyl jasmonate	5 μl MJ was dissolved in 440 μl 95% (v/v) ethanol. The solution should be freshly prepared before each experiment.
Sodium chloride	Stock solution of 25% w/v was prepared in water, sterilized by autoclaving at 121°C for 20 min, andkept at 4°C.

2.1.8.2 Effect of elicitor types used at different concentrations

A single stock culture grown in 500 ml Erlenmeyer flask at 25 °C on a rotary shaker was used as an inoculum for all experimental flasks. All experiments were carried out in 250 ml Erlenmeyer flasks containing 50 ml of MS liquid medium supplemented with 0.3 mg/l BA and 0.1 mg/l NAA and inoculated with 3 g fresh weight of 21 day old cell suspension cultures. Various types of elicitors at different concentrations (shown as follows) were added to differently old cell suspension cultures within the linear growth phase. The effect of elicitor was studied according to cell cultures harvested at various time up to 35 days after the addition of elicitors.

Elicitors	**Concentrations**			
Chitosan (mg/l)	10	20	50	100
Methyl jasmonate (μM)	50	100	200	400
Sodium chloride (%, w/v)	0.2	0.5	1	2
Yeast extract (g/l)	1	2	4	8

Three individual cultures were used for each experiment. Controlcultures were treated with distilled water (2 ml) or ethanol (400 μl).

2.2 Biochemical methods

2.2.1 Enzyme extraction

To extract proteins from *C. cajan* leaves, six commonly used buffers were selected and prepared with different pH values to cover the effective pH range of the buffers (Table 3). The salt concentration in all solutions was 100 mM and, before use, 15% (w/v) insoluble Polyclar AT (Serva, Heidelberg, Germany) was added toabsorb phenolic compounds and 10 mM dithiothreitol (DTT) as a reducing agent. PMSF (10 µM) was added to the extraction solution as an inhibitor of serine protease.
Frozen tissues were ground to fine powder under liquid nitrogen using a mortar and pestle. The powdered tissue (0.2 g) was placed in a 2-ml tube and resuspended in 1 ml of freshly prepared extraction solution. After a thorough vortexing, the mixture was incubated on ice for 10 min, during which it was vortexed a few times, and then centrifuged at 10,000 rpm at 4 °C for 10 min. The supernatant was transferred to a new 2 ml tube and extracted a second time with 1 ml of extraction solution. The extract was incubated on ice for 5 min, during which it was vortexed a few times, and then centrifuged at 10,000 rpm at 4 °C for 10 min. The supernatant (2.5 ml) was collected, and the low molecular mass substances in the supernatant were removed by gel filtration through a PD_{10} column (GE Healthcare, Freiburg, Germany), which was already equilibrated with 25 ml buffer (5 x elution volume of column), the high molecular weight fraction was eluted with 3.5 ml of buffer solution. This crude protein extracts were used in testing activities of CHS, STS and STCS.

Table 3. Solutions used for protein extraction from *C. cajan.*

Buffer	pH
Sodium citrate	4
Sodium acetate	5
HEPES	7
Sodium phosphate	7
Tris–HCl	8
Sodium borate	9

2.2.2 Determination of protein content

Colorimetric measurement of protein was carried out using the Bradford protein assay (Bradford, 1976). Coomassie Brilliant Blue G-250 interacts with aromatic and basic amino acids of the protein. The negative charged sulfonate anions of the dye give charges to cationic amino acids, which results in exposure of the hydrophobic part of the protein which interacts by the phenyl residues of the dye by van der Waals forces. Upon interaction with protein, the absorption maximum of the dye shifts from 465 nm (red, free form) to 595 nm (blue, bound form). A standard curve is constructed using serial dilutions of bovine serum albumin solution (1 mg/ml). Measured samples consisted of 900 µl Bradford solution completed to 1 ml with the protein solution and the buffer. The blue color reached maximum after two minutes and is stable for 1 h.

2.2.3 Enzyme assays for stilbenecarboxylate synthase (STCS) activity

The standard assay was carried out in a final volume of 250 µl containing 26 µM starter CoA-ester (Cinnamoyl-CoA, *p*-coumaroyl-CoA, dihydro-cinnamoyl-CoA and dihydro-*p*-coumaroyl-CoA), 56µM malonyl-CoA asextender molecule and 0.1 M KH_2PO_4 buffer (pH 7.0). Protein amount was 50 µg crude extract or 6-12 µg pure recombinant protein. The reaction was incubated at 37°C for 30 min. The reaction was stopped with 25 µl 10% HCl and the products were extracted twice with 250 µl ethylacetate by vigorous mixing for 1 min and centrifugation at 13,000 rpm for 10 min. The organic phases were dried under vacuum and the residue was dissolved in 50 µl of methanol. Analysis of the enzymatic products was performed by HPLC.

2.2.4 Mixed enzyme assay with type III PKSs and polyketide cyclases

Mixed enzyme activity was assayed essentially as described by Page et al. (2012). Enzyme assays were performed in 250 µL with 0.1 M KH_2PO_4 buffer (pH 7.0), 26 µM starter CoA ester and 56 µM malonyl-CoA as extender molecule. Purified recombinant PKSs (PASs 10 µg) and polyketide cyclases (DAC 7 µg, BEC 7 µg, and CHC 7 µg) were added in combination. Reaction mixtures were incubated at 30 °C for 60 min. Reactions were stopped with 25 µl 10% HCl and the products were extracted twice with 250 µl ethyl acetate. The organic layer was combined and dried under vacuum. The residues were dissolved in 50 µl of methanol for HPLC analysis.

2.2.5 HPLC analysis of enzyme assays

Analysis of enzymatic products was performed by HPLC (Agilent 1200 series) using a Hyper Clone ODS column (C18, 150 x 4.6 mm, 5µm) with flow rate 0.5 ml min^{-1}. The mobile phase consisted of water containing 0.1 % formic acid (A) and methanol (B) with step gradient elution as follows:

Gradient		**Detection wavelengths (nm)**
Time (min)	Methanol (%)	
0	40	
2	40	254
22	70	280
24	70	304
29	100	323
32	100	350
35	40	
45	40	

The identity of the enzymatic products was confirmed by UV spectroscopy, cochromatography with authentic reference compounds and by LC-MS for related products in the standard assay.

2.2.6 Electrospray ionization-mass spectrometry (ESI-MS)

Some purified products from enzyme assays (0.1 mM) were analyzed and characterized *via* ESI MS/MS. The purified solution was analyzed by direct infusion with a flow rate of 5-10 µl/min. Mass spectrum analysis was carried out through "Positive Full-Scan Mode". MS/MS experiments were done with EPI^+ (enhanced product ion scan, positive mode). Instrument tuning and mass calibration were based on the molecular ion peak $[M+H]^+$ through the whole fragmentation pattern.

2.2.7 Biochemical characterization of PASs and CHSs from *C. cajan*

For biochemical characterization of PAS and CHS enzymes, pure recombinant protein was used. All incubations were performed in triplicate and average values were calculated.

2.2.7.1 Determination of pH and temperature optima

To study the pH optimum, 100 mM potassium phosphate buffer ranging from pH 5.0-7.0, 100 mM sodium citrate buffer ranging from pH 1.0-4.5 and 100 mM Tris-HCl buffer ranging from pH 7.5-10.0 were used for optimum buffer capacity. Atthe optimum pH value, another series of incubations were performed at different temperatures between 20-60°C

2.2.7.2 Linearity with protein amount and incubation time

The incubations were carried out as in (II.2.2.3) at the pH and temperature optima. The amount of enzymatically formed products was determined as a function of the protein amount in the standard assay (0.25 -30 µg) and incubation time (3-60 min).

2.2.7.3 Study of substrate specificity

At the pH, temperature, protein amount and incubation time optima, enzyme assays were performed using malonyl-CoA as extender substrate and a series of starter substrates: cinnamoyl-CoA, *p*-coumaroyl-CoA, dihydro-*p*-coumaroyl-CoA, dihydro-cinnamoyl-CoA, acetyl-CoA, benzoyl-CoA, *m*-coumaroyl-CoA, *o*-coumaroyl-CoA and salicoyl-CoA. After incubation, the products were extracted and analyzed by HPLC. Three independent experiments were performed and mean values were calculated by radiodetector-coupled HPLC.

2.2.7.4 Determination of kinetic parameters

At optimum pH and temperature, each assay containing 2µg proteins, the kinetic properties were calculated from Lineweaver-Burk plots. Appropriate incubation time (10 min) was chosen so that the reaction velocity was linear during the assay period.

In case of PAS, for cinnamoyl-CoA the concentration was between 1-72 µM and malonyl-CoA was kept constant at 56 µM. For malonyl-CoA, the concentration was varied between 6-200 µM while cinnamoyl-CoA was constant at 26 µM. For *p*-coumaroyl-CoA, the concentration was changed between 1-96 µM and malonyl-CoA was kept constant at 56 µM. For malonyl-CoA, the concentration was varied between 6-200 µM while *p*-coumaroyl-CoA was constant at 32 µM.

In case of C.CHSs, the concentration of cinnamoyl-CoA was between 1-64 µM and malonyl-CoA was keptconstant at 78 µM. For malonyl-CoA, the concentration was varied between 6-200 µM while cinnamoyl-CoA was constant at 30 µM. For *p*-coumaroyl-CoA, the concentration was changed between 1-64 µM and malonyl-CoA was kept constant at 78 µM. For malonyl-CoA, the concentration was varied between 6-200 µM while *p*-coumaroyl-CoA was constant at 27 µM.

2.3 Molecular biology methods

2.3.1 Isolation of nucleic acids

2.3.1.1 Isolation of total RNA

The RNeasy Plant Mini Kit (Qiagen) was used for the isolation of total RNA from different plant organs. Freshly harvested plant samples (100 mg) or frozen (-80°C) samples (100 mg) were homogenized with liquid nitrogen and disrupted immediately in a lysis buffer containing a strong denaturating agent. Following cell lysis, centrifugation through QIAshredder ^{TM} was conducted to remove insoluble materiall and to reduce the viscosity of the lysate. The cleared lysate was transferred to RNeasy Mini Spin columns in the presence of absolute ethanol, which promotes aselective binding of RNA to the silica-gel-based membrane of the RNeasy Mini Spincolumn. By centrifugation the contaminants including small RNAs were efficientlywashed away, high quality RNAs were eluted in RNase-free water and stored at -80°C until used. The RNA concentration was determined photometriclly as described in 3.2.

2.3.1.2 Isolation of Genomic DNA

The DNeasy Plant Mini Kit (Qiagen) was used for the isolation of genomic DNA from different plant organs. Plant material (100 mg) was pulverized in liquid nitrogen and lysed in a buffer containing RNase. Following lysis, cell debris and precipitates were removed by centrifugation through a QIAshreder TM column. Binding buffer and ethanol were added to the clear lysate to allow DNA adsorption on the membrane in presence of high concentration of chaotropic salts, which removed water from hydrated molecules in solution, whereas contaminants such as proteins and polysaccharides were efficiently removed by two wash steps. Pure DNA was elutedin 10 mM Tris-HCL buffer pH 8.0.

2.3.2 Determination of RNA and DNA concentration

The concentration of nucleic acids was determined by measuring the absorbance value of the DNA or RNA samples at a wavelength of 260 nm. One unit absorbance value at wavelength 260 nm corresponds to 40 ng/µl single-stranded RNA or 50 ng/µl double-stranded DNA (Sambrook et al., 2001).

Concentration of RNA = absorbance at 260 nm×dilution factor×40 µg/ml.

Concentration of DNA = absorbance at 260 nm×dilution factor×50 µg/ml.

Purity is determined by calculating the ratio of absorbance at 260 nm to absorbance at

280nm. Pure samples having an A_{260}/A_{280} ratio of 1.7-1.9 ensure the appropriate quality. Contaminations with proteins or phenolics would reduce this value due to their absorbance at 280 nm.

2.3.3 Reverse transcription (RT)

A reverse transcriptase enzyme is known as RNA-dependent DNA polymerase which is used to synthesize the first strand cDNA (cDNA is a strand of DNA complementary to appropriate mRNA) from total RNA, using an oligo-dT primer which annealed to the poly A tail of the mRNA. Reverse transcription of total RNA is carried out using RevertAid™ H minus reverse transcriptase. It is genetically engineered to demolish the RNase H activity, so it will not degrade RNA-DNA hybrids, which in turn will yield higher yields of full-length cDNA from long templates. RevertAid™ H minus reverse transcriptase maintains activity over a wide temperature range (42-55°C) and is capable of full-length first strand cDNA synthesis.The procedure consists of two stages: 1. denaturing the RNA to get rid of the secondary structure; 2. the reverse transcription reaction. The reverse transcription reaction product can be directly used or stored at -20°C.

cDNA synthesis reaction

Component	**Volume**
1µg total RNA	x µl
Oligo dT primer (5'CDS) 10 pmol	1 µl
Nuclease-free water up to	11 µl
Denaturation of RNA at 70°C for 5 min followed by briefly cooling on ice and addition of the following reagents	
5x reaction buffer	4 µl
10 mM dNTP mix	2 µl
RNase inhibitor (40 U/µl)	0.5 µl
The mixture was prewarmed at 37°C for 5 min followed by addition of	
RevertAid™ H minus reverse transcriptase	1 µl
Nuclease-free water up to	20 µl
The reaction mixture was incubated at 42°C for 90 min, followed by heat inactivation of the enzyme at 70°C for 15 min.	

2.3.4 Design of gene-specific primers

A gene-specific primer is designed to have a sequence which is the reverse complement of a region of template or target DNA to which the primer anneals. As mentioned elsewhere, certain rules should be taken in consideration by primer design. It is not necessary that the designed primers comply with all the mentioned recommendations.

1. Primer length is between 15-30 nucleotides.
2. GC content is between 40-60%
3. Long stretches of G or C should be avoided.
4. Primers should end with G or C.
5. Difference in Tm between forward and reverse primers should not exceed 1°C.
6. Complementarity between the primer pairs should be avoided to prevent primer

dimer formation, and complementarity within the primer should be avoided to prevent hairpin formation.

7. Primers should be checked against their template DNA for undesirable complementarity.
8. The annealing temperature must be approximately 5°C lower than the melting temperature.

The portion of template DNA to be amplified is selected. Forward as well as reverse primers are designed as mentioned above. The length of primer is adjusted to obtain a high T_m. Consequently, PCR reaction at high annealing temperature can be carried out. The higher the annealing temperature is, the higher the specificity of the reaction. The lower the annealing temperature is, the higher the chance for miss-priming and wrong amplification.
In course of this study, primers were designed as follows:
Candidate sequences from the *C. cajan* genome sequence were aligned; the ones with a high degree of similarity are grouped together. Primer pairs aiming at amplifying the full lengths of these candidates are designed, taking in consideration the factors mentioned above. Then restriction sites are introduced at beginning of the primers. Between the start codon of the forward primer and the introduced restriction site, one, two, or no nucleotide should be introduced to keep the coding sequence in frame. This depends on the nucleotide sequence of the vector and the restriction site introduced.

2.3.5 Polymerase Chain Reaction (PCR)

In the last decades, PCR has rapidly evolved in the molecular biology field. PCR is aprimer-mediated enzymatic amplification of specifically cloned or genomic DNA sequences. The “chain reaction" of the technique refers to an increase in the amount of target DNA obtained through successive cycles of amplification. Three cycles are applied: a) denaturation: at 95°C to separate double stranded DNA into single strands. b) Annealing: to allow binding of primers to single DNA strands, this binding is the first starting point for DNA polymerase enzyme. The annealing temperature depends on melting temperature (T_m) of primers. It should be neither so high (no binding occurs) nor so low (miss-priming occurs). Usually it is about 5°C less than the Tm of primers. c) extension or elongation: DNA polymerase extends the primer till getting double stranded DNA again. This step is carried out at 72°C, the best temperature for DNA polymerase activity. These sequences of cycles will double the initial amount of DNA template. After 20 cycles, one gets one million double of the initial amount of initial template DNA, so traces of DNA can be detected and amplified to measurable amounts using PCR. A variety of recombinant DNA polymerases are available on the market. Taq-DNA polymerase was used in semi-quantitative RT-PCR study. However, this enzyme has a high error rate (10^{-5}). Therefore, high-fidelity Phusion DNA polymerase was used in amplification of clones, which will be downstream used in protein expression. This enzyme possesses 3`-5`-exonuclease activity which allows the proof reading activity with an error rate of 10^{-7}. The composition of each reaction as well as the temperature program is slightly different and by using Phusion DNA polymerase the order of adding the reaction components is important.

Standard PCR using Taq DNA	Using Phusion DNA polymerase (20 µl

polymerase (25 µl reaction)		reaction)	
Template DNA	1 µl	Template DNA	1 µl
Forward primer (10 pmol)	1 µl	Forward primer (10 pmol)	1 µl
Reverse primer (10 pmol)	1 µl	Reverse primer (10 pmol)	1 µl
10× reaction buffer	2.5 µl	5×phusion buffer	4 µl
dNTP (10 mM each)	1 µl	dNTP (10 mM each)	0.4 µl
Enzyme (5 U/µl)	1 µl	Enzyme (2U/µl)	0.4 µl
PCR Water up to	25 µl	PCR Water up to	20 µl

PCR program

	Taq DNA polymerase		Phusion DNA polymerase	
Step	Temp. (°C)	Time	Temp. (°C)	Time
Initial denaturation	95	3 min	98	30 s
Denatruation	95	30 s	98	10 s
Annealing	50	30 s	50	45 s
Elongation	72	3 min (1min/kb)	72	90 s (30 s/kb)
Final extension	72	10 min	72	10 min

2.3.5.1 Touch down PCR

Touch down PCR uses an annealing temperature during the initial PCR cycles that is 5-10°C higher than melting temperature (T_m) of the primers. In subsequent cycles, theannealing temperature is decreased in increments of 0.5°C per cycle until the temperature was reached 5°C below the Tm of the primers.

Touch down PCR programme

Step	**Temp [°C]**	**Time [sec]**	**Cycles**
1	95	180	
2	70	Pause	
3	95	45	
4	T_m+5°C	45	Tm decrease 0.5 degree each cycle
5	72	60	10x from step 3
6	95	45	
7	T_m– 5°C	45	
8	72	120	30x from step 6
9	72	600	
10	15	Pause	

2.3.6 Agarose gel electrophoresis

It is used for analysis and separation of DNA and RNA based on their size and charge. Small DNA molecules migrate faster than larger molecules through the agarose matrix

under effect of electric current. Based on the size range of DNA molecules in the sample, agarose gel is prepared in the range of 0.5-2% agarose in TAE buffer. The higher the proportions of agarose, the smaller the pore size of the matrix and the slower the migration rate. The mixture was boiled in a microwave oven to dissolve the agarose. After cooling down the mixture to 60°C, ethidium bromide (0.5 µg/ml) was added and the solution was poured into the gel tray. Ethidium is a flat molecule, just the right size to get in between the stacked base-pairs of double-stranded DNA. This, and its fluorescence, is what make it a sensitive dye for detecting DNA. Since ethidium bromide is acarcinogen, all steps were performed under the fume hood. Samples were loaded on gel after mixing with loading dye along with DNA or RNA ladder. Then, samples were run in a gel chamber filled with TAE buffer and under electric current (110 V). The gel was developed till the visible dyes in the loaded sample migrate 2/3 the distance. The gel is visualized under UV-equipped transilluminator, and photographed with multiimage™ light cabinet (Biozyme).

2.3.7 Extraction and purification of DNA from agarose gels

After gel electrophoresis, DNA fragments had to be eluted from gel for further applications such as restriction enzyme digestion and/or ligation. Purification of DNA fragments was performed by using the Innu PREP DOUBLE Pure Kit (Analytic Jena biosolution) to obtain clean DNA fragment. The right size DNA band was excised under UV light, the DNA fragment was adsorbed to a silica gel membrane in the presence of a high concentration of chaotropic salts, the binding mixture was loaded directly into a Nucleospin column, contaminants like salts and soluble macromolecular components were removed away by washing and the DNA was eluted with water or elution buffer.

2.3.8 Digestion of PCR products or vectors

For heterologous expression of a protein, the amplified ORF should be inserted into an expression vector (plasmid), which can be transferred later to the host bacteria. To achieve this, both the vector and the PCR product are digested with endonuclease(s) to produce either sticky or blunt ends. These complementary ends are then ligated. The used endonuclaease are the ones, which have no recognition sites in the target insert sequence but have in the MCS (multiple cloning sites) of the vector. In our case, the overexpression vector was pRSET B (Invitrogen). The ORFs of PASs, CHSs and cyclases were amplified with Phusion Hot start II and overexpression primers (II.1.5.2). The PCR products are digested with endonucleases and then ligated with digested dephosphorylated vector. Double digestion precautions and recommendation were determinedfrom the website: http://www.thermoscientificbio.com/webtools/doubledigest/.

The following reaction mixture was used as a standard restriction digest protocol

Component	**Volume (insert)**	**Volume (vector)**
DNA content	0.2 µg	1 µg
Endonuclease	0.5 µl	0.5 µl
10x reaction buffer	2 µl	2 µl
Nuclease free water up to	20 µl	20 µl
Reaction was incubated at 37°C for 2 h		

Digestion products were subjected to purification using a kit to exclude residual protein and salts which may inhibit the ligation reaction. The digested vector is subjected to dephosphorylation using shrimp alkaline phosphatase. Dephosphorylation of the 5` group of the vector ends prevent self-ligation of the vector during the ligation reaction. The composition of the dephosphrylation reaction was as follows.

Vector solution	10-40 µl
10x reaction buffer	2 µl
SAP enzyme 1 U/1 µl	1 µl
Water ad	50 µl

The reaction was incubated for 15 min at 37°C, and then the enzyme is inactivated by 65°C for 15 min. The reaction product is purified in order to get rid of residual proteins and components which can inhibit the ligation product.

2.3.9 Ligation of DNA fragments

To promote the ligation reaction, the insert amount should be 2-6 fold more than the vector amount. The vector and insert are mixed together in PCR tubes and kept at 55°C for 5 min, then chilled on ice. This step is done to avoid mispriming of the sticky ends. The tubes are centrifuged and then the buffer and the enzyme are added and the reaction is incubated by 16°C overnight. A negative control reaction which contains all the components except the insert is done in parallel. Ligation reaction (5 µl) is used to transform 50 µl competent cells.

Component	**Volume (µl)**
Digested vector	1
Digested insert	6
10x ligation buffer	1
T4-DNA-Ligase (5U/ µl)	0.5
Nuclease free water up to	10

2.3.10 Transfer the constructed plasmid DNA into *E.coli*

Chemically competent cells are prepared in our laboratory by the calcium chloride method (Cohen et al., 1972; Dagert and Ehrlich, 1979; Mandel and Higa, 1992). This treatment enhances the attachment of plasmid DNA to the bacterial membrane. Transformations occurred into *E. coli* DH5α and BL21 (DE3) pLysS (Ausubel et al., 1994).

2.3.10.1 Transformation into DH5α

The *E. coli* strain DH5α was used for initial cloning of target DNA in pRSET B vector and also for plasmid maintenance because this strain has high transformation efficiency and high plasmid yields. Ligation product (5 µl) is added to competent cells (50 µl) and

incubated for 25 min on ice, then transferred to a water bath at 42°C for 50 s followed by immediate incubation on ice for 2 min. SOC medium (250 µl) was added and bacterial suspension is incubated by 37 °C for 1 hour. The whole bacterial suspension is plated on LB-agar plates containing ampicillin (100 µg/ml).

2.3.10.2 Transformation into BL 21 (DE3) pLysS

Plasmid containing DNA for heterologous protein expression was transformed into the expression host *E. coli* BL21 cells. BL 21 is the strain of choice for high-level geneexpression and production of recombinant protein in a bacterial system. The transformation procedure was the same as with DH5α, except that the heat shock time was for 20 seconds. After one hour ofincubation at 37°C, about 80-100 µl of bacterial suspension was finally plated on LB-agar medium containing ampicillin (100 µg/ml) and chloramphenicol (60 µg/ml). Ampicillin resistance is acquired from the expression vector and chloramphenicol resistance is acquired from the host cells. Bacterial cultures containing plasmids with target insert were stored at -80°C in autoclaved solution of 20% glycerol/LB-medium. This continuous cultures from *E. coli* BL21 (DE3) containing the desired gene can be used as a permanent supply to start protein expression.

2.3.11 Plasmid isolation

It was carried out as described by Birnboim and Doly (1979). A single colony of the transformed DH5α was inoculated into 5 ml LB medium containing 20 µg/ml ampicillin and grown over night at 37°C. On the following day, 4 ml cultures were centrifuged. Bacterial pellets were reuspended in ice-cooled buffer I (300 µl) containing RNase A, then buffer II (300 µl) was added and the bacterial suspension was inverted cautiously 6 times and incubate at room termperature not longer than 5 min. Lysis of the cell wall took place in addition to denaturation of large chromosomal DNA. RNA is destroyed by RNase A. Precipitation of proteins and denaturation of large chromosomal DNA were done by adding buffer III (300 µl), cautious inversion (6) and incubation on ice for 20 min. Centrifugation at 13,000 rpm for 10 min was done to exclude the denatured proteins and DNA. The supernatant containing the plasmid DNA (800 µl) was transferred to a new Eppendorf tube. Residual contaminants and hydrolysed protein were extracted by vortexing with 800 µl chloroform followed by centrifugation at 13,000 rpm for 10 min. The aqueous layer was transferred to a new Eppendorf tube. Isopropanol (0.7 volume) was added, vortexed and followed by centrifugation at 13,000 rpm for 10 min to precipitate plasmid DNA. The pellets were washed with 70% ethanol (500 µl) followed by centrifugation at 13,000 rpm for 10 min. The supernatant was discarded and the plasmid pellets were dried by 37°C and then dissolved in 50 µl water. Determination of plasmid concentration is done as mentioned under II.2.3.2. Restriction analysis to test for the presence of the expected insert is carried out as mentioned under II.2.3.8.

2.3.12 Heterologous expression of recombinant protein in *E. coli*

As mentioned before, selection for *E. coli* BL21 bacteria containing the desired clone were carried out by growing the transformed *E. coli* BL21 on agar plates containing ampicillin and chloramphenicol. Ten ml LB medium containing ampicillin (100 µg/ml) and

chloramphenicol (60 μg/ml) was inoculated with one colony and grown overnight (18 h) by 37 °C. On the following day, 4 ml from the overnight culture were transferred to 100 ml LB culture containing ampicillin (100 μg/ml). The cultures were grown by 37°C until OD600 of 0.6-0.8 was reached. Freshly prepared IPTG was added at a final concentration of 0.5 mM in order to induce the production of the recombinant protein. The cultures were then incubated at 23°C for 6 h, and then the cells were harvested by centrifugation and could be kept at -20°C for the next step. Protein expression was performed by low temperature to slow the metabolic processing of the bacteria and then giving the opportunity for the protein to be rightly folded.

2.3.13 Extraction and purification of expressed proteins from *E. coli* cells

Heterologously expressed protein was extracted from *E. coli* cells to determine its biochemical activity. Mechanical disruption of the cell membrane (sonication) is a common method for breaking the cell wall and isolating the soluble protein. All procedures were carried out at 4°C. Lysis buffer (5 ml) was added to the bacterial pellets. Cell wall was disrupted by sonifier for 5 min at duty cycle 40% and output control of 1.5. Cell debris was separated by centrifugation. The purification procedure was performed on a nickel-nitrilotriacetic acid (Ni-NTA) protein purification system. Ni-NTA slurry (200 μl) was added to the supernatant. After shaking at 4°C for 1 h, the mixture was loaded into acolumn. Affinity purification started with 5 ml washing buffer. The His_6-tagged-fusion protein was eluted using 2.5 ml elution buffer. Imidazole (used for elution) was removed from the eluate by gel filtration through a PD10 column equilibrated with 0.1 M Tris-HCl buffer pH 7.5.

2.3.14 SDS-PAGE gel electrophoresis

It is used to separate proteins according to their masses. The protein samples are denaturated in presence of the sample buffer at 95°C for 5 min. During denaturation, the protein is defolded and becomes uniformly coated with negative charges from SDS in the loading buffer. This process will cancel the differences between proteins regarding charge. Glyceol in the sample buffer ensures that samples will sink and settle down in the pockets of the SDS-PAGE gel, while bromophenol blue dye will help in tracing the development of the gel. Bromophenol blue is a small size dye whose particles migrate faster than any protein. When samples are loaded on the SDS-PAGE gel, proteins will migrate through the two zones; namely, stacking zone gel and resolving zone gel. In the stacking zone, proteins migrate through large pores; no separation takes place, rather all the proteins in a sample will collect as a thin compact band at front of the resolving gel. The resolving gel has a narrower pore size, and proteins will be separated according to their masses. Low molecular weight proteins will run faster than large molecular weight proteins. Bisacrylamide interconnect the polymer chains of acrylamide. The concentration of acrylamide and bisacrylamide in the separating gel was 12%, which allowed the highest resolution of proteins between 10 and 200 kDa. The polymerization reaction is initiated by addition of the ammonium peroxydisulfate (APS), a radical initiator. TEMED, N,N,N′,N′-tetramethylethylenediamine, is a polymerization catalyst. The degree of polymerization, and hence the pore size of the gel, is dependent on the amount of acrylamide and the pH of polymerization reaction. The resolving gel ingredients are mixed

together in a flask, then poured between the glass plates of the instrument and overlaid with water to avoid interaction with air. After gel formation, water is absorbed away with filter paper. The same procedures are carried out for the stacking gel, except that a comb is added to create the sample pockets, no water is used. Protein samples are prepared as mentioned before. Bacterial pellets (from 1 ml bacterial suspension) are suspended in sample buffer (100 µl) and denatured in the same way as protein samples. To estimate the molecular mass of the separated protein, a 10-170 kDa protein marker was loaded in parallel with the samples. Samples are loaded on the gel and run at 200v. The run ends when the dye bromophenol blue reaches the front of the gel. The gel is taken out for staining. The gel is stained by immersing in coomassie blue solution under gentle shaking for 30 min and then destained by immersing in destaining solution overnight.

2.3.15 Databases and software

- The Genome Database for *C. cajan*: http://www.icrisat.org/gt-bt/iipg/Home.html

It is a curated and integrated web-based relational database, providing centralized access to *C. cajan* genomics and genetics data and analysis tools to facilitate cross-species utilization of data. It was used to search the *C. cajan* genome for possible STCSs candidates.

- Basic Local Alignment Search Tool (BLAST): http://blast.ncbi.nlm.nih.gov/Blast.cg

It uses nucleotide or protein query to search nucleotide and protein databases available by website or any set of sequences supplied by the user.

- DNA star lasergene

A software package consists of different programs to manipulate and analyze DNA and protein sequences. From these programs, two were used.

MegAlign generates pairwise and multiple sequence alignments of DNA or protein or a combination of the two quickly and accurately. It was used in alignment of different sequences in addition to calculating the percent identity between different sequences.

EditSeq is a sequence editor and import/export tool. It helped by selecting regions for primer design and locating different nucleotides.

- Mega 5 (Molecular Evolutionary Genetic Analysis) was used to infer the phylogenetic tree (Tamura et al., 2011).

III. Results

1. Establishment of callus and cell suspension cultures of *C. cajan*

1.1 Germination of *C. cajan* seeds

The best sterilization procedures were achieved when 50% chlorox (outside the laminar flow cabinet for 1 min) and 70% ethanol (inside the laminar flow cabinet for 1 min) were utilized. Surface-sterilized seeds cultured directly on MS basalagar medium without any hormones in the phyta jars showed 40-50% germination within 4 days, and 95-100 % germination was observed after 7 days (Fig. 19). Pre-soaked seeds resulted in better germination compared to unsoaked seeds. Seven-day-old *in vitro*-grown seedlings, which had produced the first two fully expanded simple leaves and another well-developed trifoliate leaf besides the shoot tip were used as the source of leaf explants for initiating callus cultures. Some of this kind of primary seedlings were transferred to 300 ml Erlenmeyer flasks containing the same medium without plant hormones for 8 weeks. The resulting explant sources, such as leaf, stem and root segments, were derived from these aseptic seedlings and used in the following works.

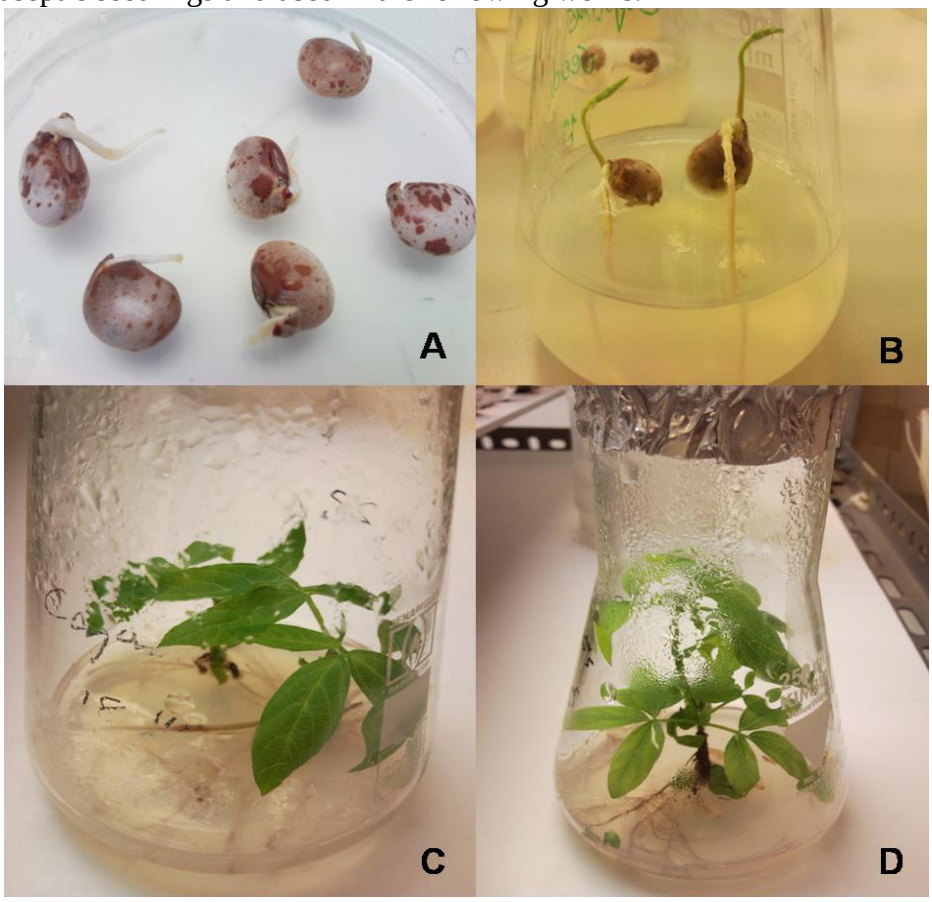


Fig. 19. Growth of *C. cajan in vitro* plants after germination. A: 4-day-old germinating seeds. B: 7-day-old seedlings. C: *C. cajan in vitro* plants (14 days old). D: *C. cajan in vitro* plants (8 weeks old).

1.2 Callus induction from leaf explants

Callus induction was observed on MS medium containing different concentrations and combinations of auxins (NAA and 2,4-D) and cytokinins (kinetin and BAP). Most of the leaf explants showed bigger growth in size only after 11 days. Callus initiation occurred from the cut ends of the explants during 3 weeks of culture in most of the leaf explants. There were generally various types of callus such as compact, bubbly and friable. The colors of callus varied from green, cream, white, brown to light green. The induction percentage, texture and color of callus were different on different media. The best hormone concentrations were chosen based on the highest percentage and quality of callus. Results of callus induction from leaf explants are shown in Tables 4 and 5.

Table 4. Effect of BAP and NAA combinations on callus induction from leaf explants.

Hormones (mg/L)		No. of explants inoculated (21 days)	Percentage of callus induction (%)	Characteristics of callus (color and texture)
NAA	BAP			
0	1.0	60	96	Light green, friable
0	1.5	60	90	Yellow, bubbly
0.1	0.3	60	100	Green, friable
0.1	0.5	60	100	Light green, friable
0.1	1.0	60	91	Dark green, friable
0.3	0.1	60	85	Green, compact
0.3	0.3	60	90	Green, compact
0.5	0	60	35	Dark green, compact
0.5	0.1	60	81	Light green, compact
0.5	0.5	60	78	Green, friable
0.5	1.0	60	97	Yellow, friable
0.5	1.5	60	92	Dark green, bubbly
0.7	0.5	60	75	Light green, compact
1.0	0	60	13	Brown, compact
1.0	0.3	60	84	Green, compact
1.0	0.5	60	78	Green, friable
1.0	0.1	60	87	Dark green, compact
1.5	0.3	60	90	Brown, compact
1.5	0.5	60	93	Dark green, friable

Table 5. Effect of 2,4-D and KT combinations on callus induction from leaf explants.

Hormones (mg/L)		No. of explants inoculated (21days)	Percentage of callus induction (%)	Characteristics of callus (color and texture)
2,4-D	KT			
0	1.0	60	22	Yellow, bubbly
0	1.5	60	17	White, bubbly
0.1	0.3	60	87	Light green, compact
0.1	0.5	60	80	Light green, friable
0.1	1.0	60	61	Light green, bubbly
0.3	0.1	60	87	Green, compact
0.3	0.3	60	90	Green, compact
0.5	0	60	97	Green, compact
0.5	0.1	60	100	Light green, compact
0.5	0.5	60	97	Green, friable
0.5	1.0	60	75	Light green, friable
0.5	1.5	60	82	Yellow, bubbly
1.0	0	60	95	Green, compact
1.0	0.3	60	85	Green, compact
1.0	0.5	60	80	Green, friable
1.0	0.1	60	86	Dark green, compact
1.5	0.3	60	90	Brown, friable
1.5	0.5	60	93	Dark green, friable

Callus induction was noticed in all media formulations but there was a wide range of variation in percentage of callus formation and the quality of callus. Although the same percentage of callus formation was found with the concentrations of 0.5 mg/L BAP and 0.1 mg/L NAA, the concentrations of 0.3 mg/L BAP and 0.1 mg/L NAA were the first to induce callus production after 22 days at a 100% rate and better quality. In the combination of 2,4-D and KT, the MS media with 0.1 mg/L KT and 0.5 mg/L 2,4-D also had the highest callus production after 21 days (100%). However, the compact callus was not good for establishment of cell suspension cultures. In combination of BAP and NAA, a remarkable increase in percentage of callus formation was obtained (0.3 mg/L BAP and 0.1 mg/L NAA) and cells were green and friable (Fig. 20). Actually, in combination of BAP and NAA, a high percentage of callus formation was observed with the high concentration of BAP but the primary callus was greenish after 11 days and then grew rapidly into light green and compact callus after 21 days if the concentration of BAP was too high. In this combination (BAP, NAA), different types of color and texture were observed and most callus samples were compact and green. Different stages of cell development were observed but most of the calli were in the globular stage. In combination of 2,4-D and KT, the highest callus induction frequency was observed with 0.1 mg/L KT and 0.5 mg/L 2,4-D. Generally, 2,4-D is a very effective auxin which was widely used for callus induction. Sometimes, the media with 2,4-D alone also can lead to high callus production. In this work, 2,4-D also played an important role in callus formation because almost all higher callus induction frequencies came from the treatment with relatively high 2,4-D concentrations.

Comparison of mean production between BAP, NAA and 2,4-D, KT showed that there was no significant difference between two combinations but the callus status was higher in BAP, NAA after 21 days. In the control (that is, basal medium without hormones),

there was a little growth extension on the explants, however, no callus production was observed during the culture period. Thus, the 0.3 mg/L BAP and 0.1 mg/L NAA combination, which resulted in the highest percentage of callus formation and the best status of callus, was chosen and added to MS solid medium used for further subcultures. All experiments were repeated three times. The response of callus on different basal media supplemented with optimized hormones was also improved, the results are shown in Table 6. In summary, MS medium revealed the best properties. Therefore, the optimized conditions of callus induction from *C. cajan* leaf explants were obtained as follows. Callus proliferation was induced on MS medium supplemented with 0.3 mg/L BA and 0.1 mg/L NAA under a 16 hr photoperiod, subculturing being performed at 21-day-intervals.

Table 6. Response of callus formation on different basal media supplemented with optimized hormone levels

Medium	No. of explants inoculated (11 days)	Percentage of callus induction (%)	Characteristics of callus (color and texture)
MS	**60**	**100**	**Green, friable**
LS	**60**	**96**	**Green, brown, friable**
B5	**60**	**80**	**Green, compact**
4X	**60**	**15**	**Brown, compact**

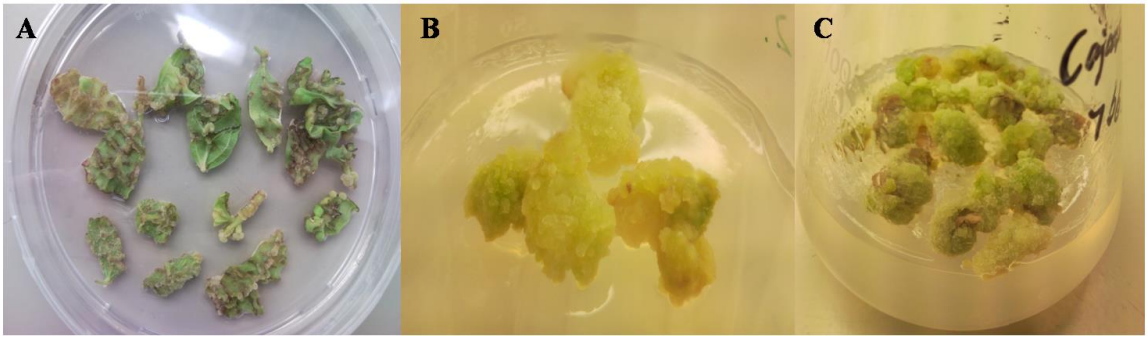


Fig. 20. Callus induction from leaf explants. A: *C. cajan* explants, B: Green and friable callus, C: Dark green and compact callus.

Growth of *C. cajan* callus under optimized conditions was characterized as follows, the biomass accumulation being expressed in gram dry weight. A typical growth curve (Fig. 21) was obtained with a lag-phase until the 3rd day of culture. After the lag phase, callus started to grow and propagated actively, entering its exponential phase of growth on the fourth day and reached its maximum on day 22. Its growth rate became negative after day 22 and the stationary phase started. Ten-day-old greenish friable callus which was in the exponential phase of growth was used to initiate cell suspension cultures.

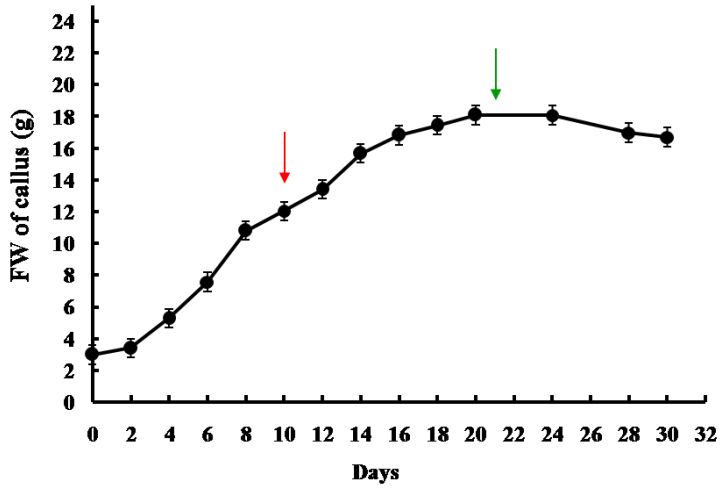


Fig. 21. Growth curve of *C. cajan* callus under optimized conditions. Red and green arrows mark callus used for cell culture initiation and start of the stationary phase, respectively.

1.3 Establishment of cell suspension cultures

Ten-day-old, leaf-derived, greenish friable callus was subcultured on MS liquid medium containing 0.3 mg/L BA and 0.1 mg/L NAA. The cultures were maintained on an incubatory shaker at 120 rpm and at 16h photoperiod and 25°C. Active division and growth of cells were observed until the 6th and 7th day of culture. The growth rate of cells was initially slow during the first 6 days of lag-phase (Fig. 22). At the start the callus showed highly vacuolated cells. In the following 6 or 7 days, cell cultures started to grow and divided actively. They entered the exponential phase of growth and reached the maximum on day 21. In this period, the cells showed two morphologically distinct shapes, namely, spherical and elongated cells. The spherical embryogenic cells under went a rapid transverse division, resulting in two-cell and four-cell structures. Later these cells divided further and resulted in the formation of a clump of cells. The proembryo divided further in multiple planes within one week. Two weeks after transferring the callus to liquid medium, some smaller pieces fell off from the bigger callus clumps. At the end of three weeks, its growth rate became negative after day 22, and a stationary phase started, so the cells were harvested from the liquid medium and transferred to fresh one. After several subcultures, a fine cell suspension culture with small cell clumps was obtained (Fig. 23).

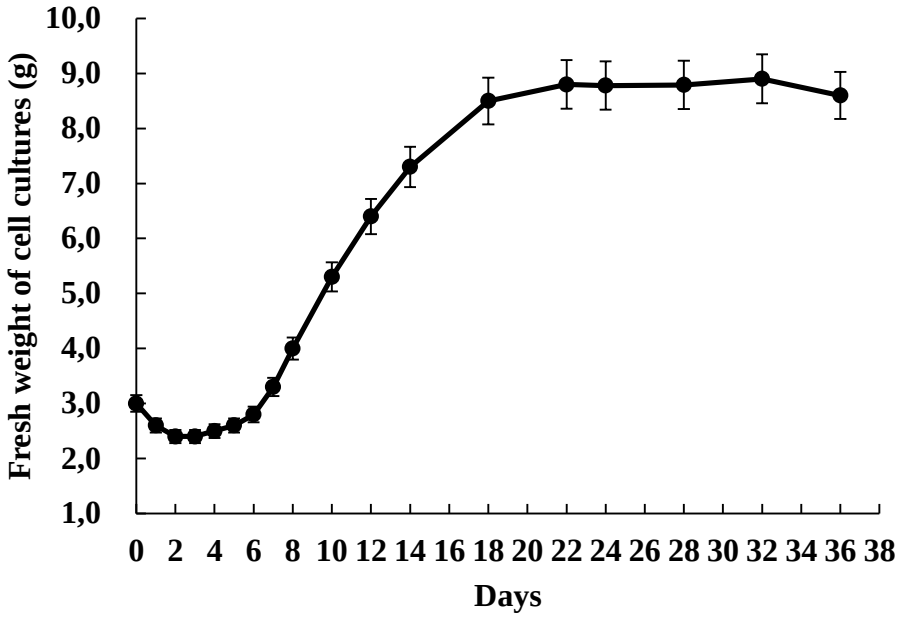


Fig. 22. Growth curve of *C. cajan* cell suspension cultures.

Fig. 23. Cell suspension culture of *C. cajan.*

In conclusion, this study established an efficient protocol for callus induction and initiation of cell suspension cultures of *C. cajan*. The long-term maintenance of callus and cell suspension cultures without an apparent change in the growth behavior was achieved for further experiments.

2. Phytochemical study of different materials of *C. cajan*

It is important to determine variations in the content of some target compounds in different materials of *C. cajan*. On the basis of these analytical informations, the start point to take samples for biochemical and molecular investigations can be accurately determined. Extraction and analysis were carried out as mentioned under II.2.1.4 and II.2.1.5 using three kinds of plant material which had been previously established. Crude extracts of the *in vitro* leaves, callus and cell suspension cultures of *C. cajan* were analyzed by HPLC. Three major compounds, pinostrobin, longistyline A and cajaninstilbene acid, were detected as target compounds (Fig. 24). The identity of these constituents was confirmed by comparison of the retention times, UV–vis spectra, and MS–MS spectra with those of authentic pinostrobin, longistyline A and cajaninstilbene acid (Fig. 25). The variation in the contents of pinostrobin, longistyline A and cajaninstilbene acid in different tissues of *C. cajan* is shown in Fig. 24. Pinostrobin could be detected clearly in all three kinds of plant material; however, there are significant differences in the content of the other two compounds. Longistyline A and cajaninstilbene acid were the quantitatively major compounds in *in vitro* leaves but less abundant in the callus and hardly detectable in cell suspension cultures.

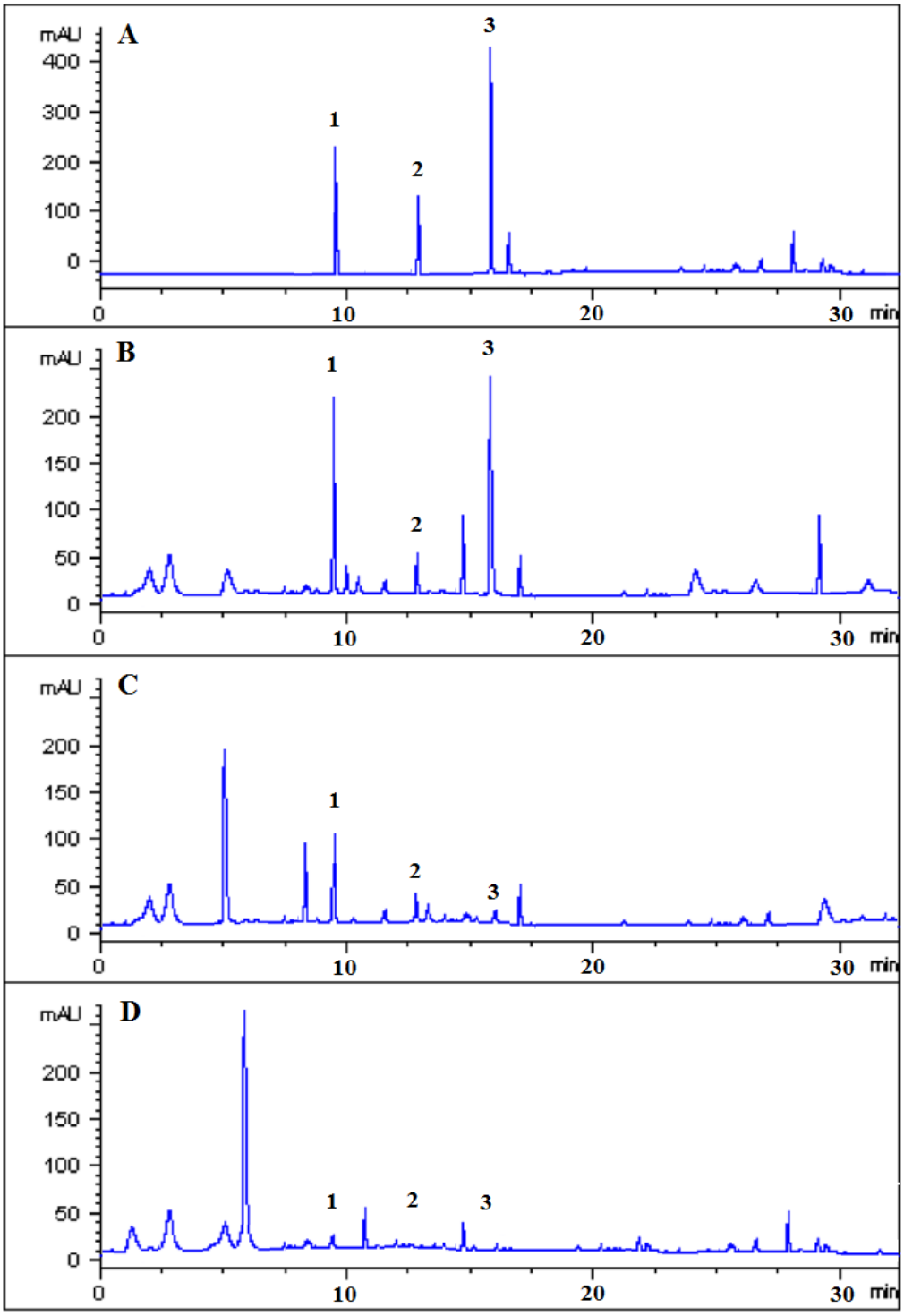


Fig. 24. Representative HPLC chromatograms at 259 nm of (A) standard mixture, (B) extracts of *in vitro* leaves, (C) extracts of callus and (D) extracts of cell suspension cultures. (1) Pinostrobin, (2) Longistyline A, and (3) Cajaninstilbene acid.

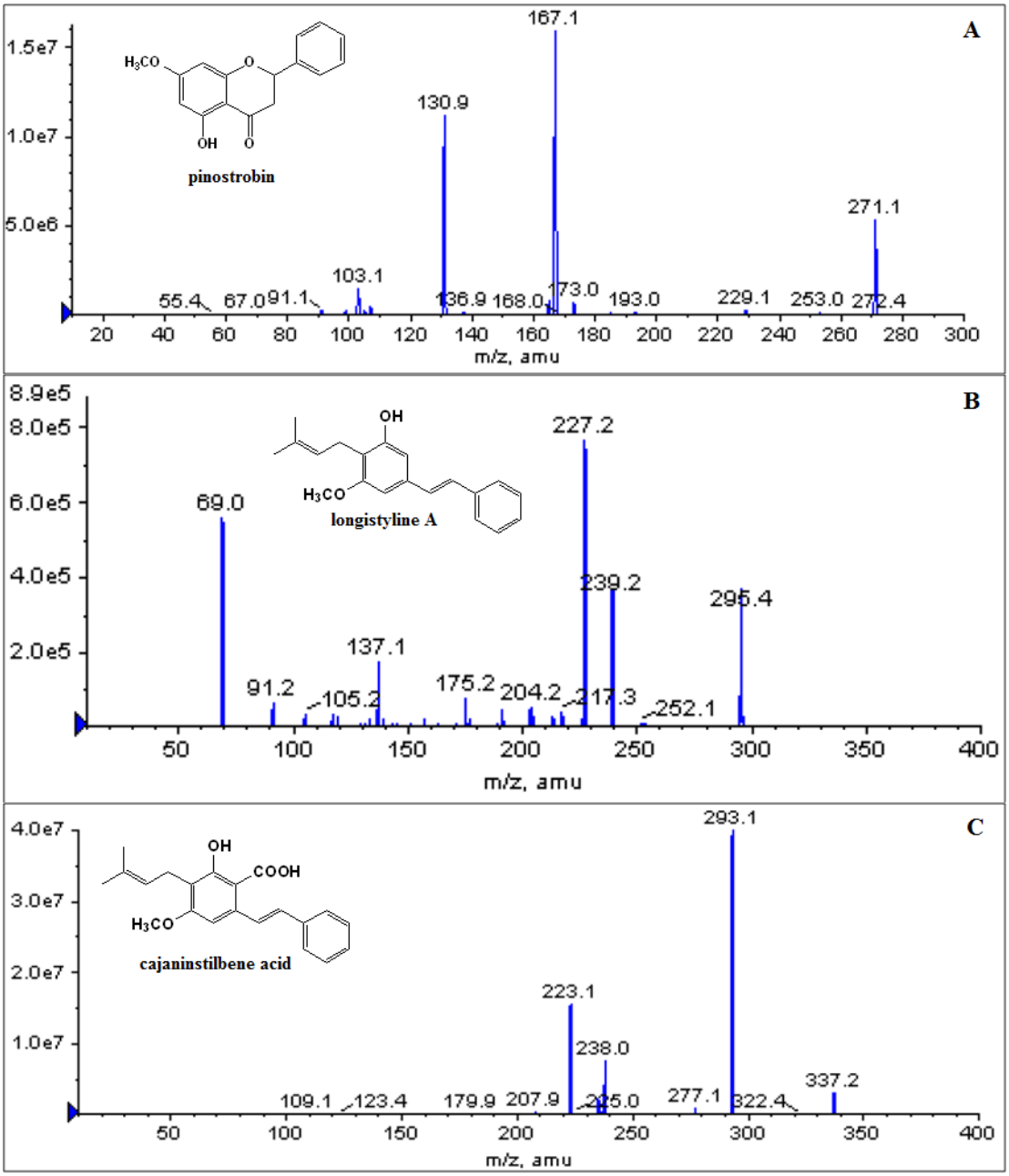


Fig. 25. Mass-spectra (LC-MS) of target compounds from different plant materials

2.1 Quantitative analysis of the target compounds in different plant materials of *C. cajan*

Quantitative analysis was carried out using HPLC. Time courses of accumulation of each compound in different materials were recorded and are shown in Fig. 26. Areas under peak values were determined. The calibration curves showed linearity for each compound, with the linear equations being summarized in Table 7. Each calibration point was carried out in triplicate**.**

Table 7. Regression data for three compounds.

Analyte	Calibration curve	r^2	Linear range (ng/mL)	LOD (ng/mL)	LOQ (ng/mL)
Pinostrobin	**y=417.5x−70.8**	**0.9973**	**200–2000**	**0.89**	**3.93**
Longistyline A	**y=497.1x−184.6**	**0.9962**	**500-2500**	**1.38**	**6.94**
Cajaninstilbene acid	**y=533.6x + 157.3**	**0.9958**	**1000-5000**	**1.62**	**7.03**

y: peak area of analyte; x: concentration of analyte (ng/mL).

All the three compounds could be detected in the leaves of *in vitro* plants at different stages, their accumulation kinetics are shown in Fig. 26. Compared to the other two analyzed components, pinostrobin showed a very different accumulation pattern. Its content increased quickly in the early stage of cultivation, reached the maximum content (2.96 mg/g DW) on day 8, and then deceased to the starting level on day 20. Although the pinostrobin concentration increased again after the 20th day and reached a second peak at day 35, this was not yet comparable to that at the 8th day. The pinostrobin content decreased again thereafter and reached the lowest content (0.87 mg/g of DW) at day 50. For longistyline A, ingeneral, the content increased with the growth of the plant. So, harvestat the beginning resulted in the lowest longistyline A content (1.27 mg/g DW). The level of longistyline A significantly increased first and then sharply decreased in the first 20 days. After that, another steady growth in content was observed and the highest value was reached at the 40th day (4.25 mg/g DW). Cajaninstilbene acid was present as the most predominant phytochemical in *C. cajan* leaves (Fig. 26) and the variation pattern was similar to that of longistyline A. The content of cajaninstilbene acid significantly changed with growth time and it seemed that there were several growth stages. In this case, the content of cajaninstilbene acid statistically significantly increased from the beginning up to 20 or 25 days, gradually increased from the 25th day, and then reached a stable level after 35 days. The higher peak values appeared at the 35th day, which were 10.45 mg/g DW.

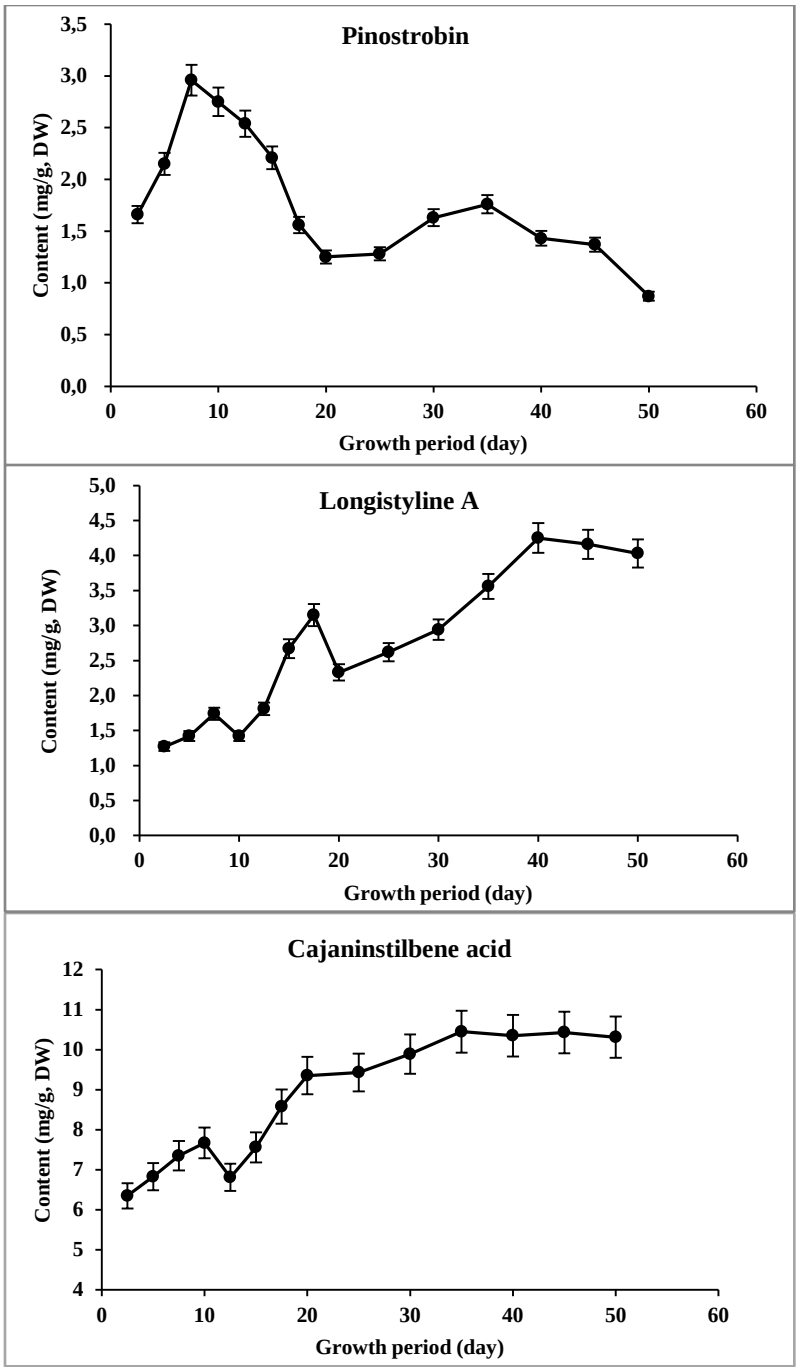


Fig. 26. Time courses of accumulation of the major compounds in *C. cajan in vitro* leaves.

In comparison to *in vitro* leaves, time courses of accumulation of compounds showed different patterns in callus (Fig. 27). Cajaninstilbene acid, which is the main compound in *in vitro* leaves, was not detectable in any growth period of callus. Although pinostrobin and longistyline A wereobserved, the amounts of the two compounds detected in the callus were much lower than those observed in *in vitro* leaves. Time courses of their

accumulation in callus are shown in Fig. 27. Measurement of pinostrobin production revealed that the maximum content (1.18 mg/g dry weight) was attained on day 12 after subculture and the content deceased toits lowest content (0.26 mg/g dry weight) on day 22. Another increase in content (0.56 mg/g dry weight) was observed on day 26 relative to the increase of fresh weight (Fig. 27). The decrease in the last period may be caused by the lack of nutrients in the medium. For longistyline A, the content significantly changed with a completely different growth pattern.The content of longistyline A significantly increased first and reached the highest values (2.02 mg/g dry weight) at the 18th day. Then, however, it sharply declined from the 20th day, followed by a relatively stable low level after day 26. These results indicated that there was an extremely intense fluctuation of the longistyline A level in *C. cajan* callus during growth.

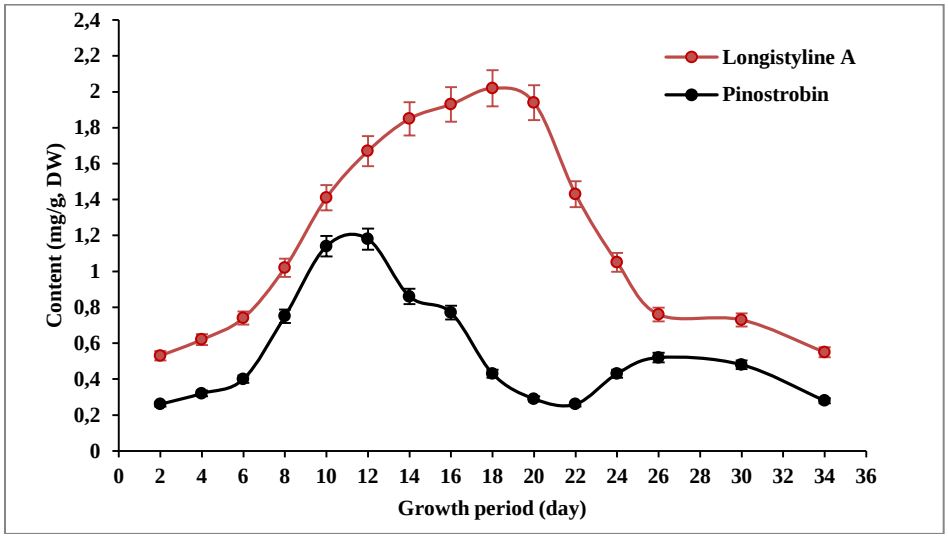


Fig. 27. Time courses of accumulation of major compounds in *C. cajan* callus.

Time courses of accumulation in *C. cajan* cell suspension culture showed an unexpected result. Almost all of the compounds were not detectable in crude extract of well stabilized cell suspension cultures. Only pinostrobin could be observed at a very low content level in the growth period of cell suspension cultures. In this case, the content of pinostrobin in cell cultures significantly changed in the range of 0 to 0.14 mg/g according to the age of the cultured cells (Fig. 28). The maximum content (0.14 mg/g dry weight) was attained on day 12 after transferring the cell suspension into fresh medium and then decreased to its lowest concentration (0.012 mg/g dry weight) on day 34. Pinostrobin was mainly present inside the cells, the amount of pinostrobin detected in the culture medium was only 0.02-0.14 % of the total amounts observed, while the other two compounds were not detected in the culture medium either.

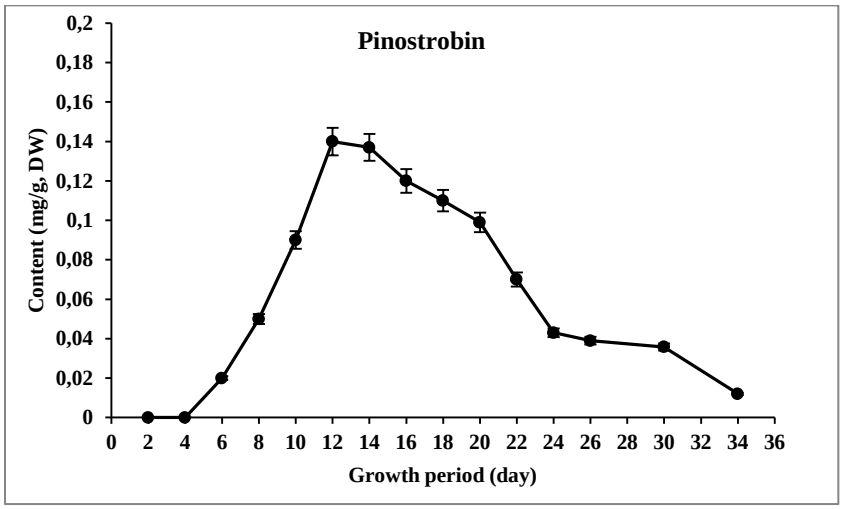


Fig. 28. Time course of accumulation of pinostrobin in *C. cajan* cell suspension culture.

In conclusion, comparison of constituents in callus, cell suspension culture and leaves of *in vitro* plant revealed that leaves of *in vitro* plant have the highest ability to form cajaninstilbene acid as a major constituent (10.45 mg/g dry weight), which is absent from callus and cell suspension culture, while pinostrobin is formed in all the three materials. The hightest contents of the three compounds were observed during the growth of leaves of *in vitro* plants.

These results indicated that there were extremely intense fluctuations of the individual compound levels in different materials during growth. It was also observed that the levels of compounds in different growth periods were dependent on the status of plant material and the culture conditions. Obviously, the biosynthesis of pinostrobin and longistyline A is more active in the leaves of *in vitro* plants than in the cultured cells, and the biosynthesis of cajaninstilbene acid is active only in the leaves of *in vitro* plants.

2.2 Variation in content of cajaninstilbene acid during growth of *C. cajan* seedlings

According to the results of the previous chapter 2.1, cajaninstilbene acid was present as the most predominant phytochemical only in leaves of *in vitro* plants without any treatment with elicitor. Therefore, further measurement ofthe onset and the course of cajaninstilbene acid accumulation after germination was carried out. Extraction and quantitative analysis of the target compound in different plant organs of *C. cajan* were carried out as mentioned above (2.1).

The variation in content of cajaninstilbene acid in *C. cajan* during the seedling growth period is shown in Fig. 29. Beginning with the emergence of *C. cajan* leaves, cajaninstilbene acid was synthesized and accumulated in leaves. Seedlings are vulnerable to pathogen attacks and accumulate chemical defenses for efficient protection early in their development, as observed for many secondary metabolites. As a whole, the levels of cajaninstilbene acid in leaves fluctuatedly increased with the growth of *C. cajan* seedlings. The details of accumulation of the target compound in leaves of *C. cajan* after

germination were shown in chapter 2.1. The variation in content of compounds correlates with the metabolic balance of plants and external environmental disturbances. Leaves are the organ which is more susceptible to environmental changes and pathogen attacks. Hence, the secondary metabolites are generally generated in high levels for efficient protection.

In *C. cajan* stems, the contents of cajaninstilbene acid increased rapidly at the early stage of seedling growth (Fig. 29). A possible reason is that at the initial seedling growth stage, the seedlings of *C. cajan* are very tender. Therefore, the stems require the same chemical defenses as the leaves. Furthermore, the secondary metabolites synthesized in the leaves may also bepartly transported into the stems with the water transportation. Later, along with the growing of *C. cajan* seedlings, the levels of cajaninstilbene acid in stems declined gradually and eventually reached steady levels during the 35 days of investigation. The highest content (2.83 mg/g DW) of cajaninstilbene acid was observed at the 15th day, which was significantly lower than that in leaves.

In *C. cajan* roots, cajaninstilbene acid was detected in the seedling growth period. Its content changed similarly to that in the stems; it firstly increased with the growth of *C. cajan* seedlings, followed by a decline and then became stable. The contents of cajaninstilbene acid were significantly lower than those in stems and leaves. The maximum was 2.05 mg/g DW, which was observed at the 20th day. However, cajaninstilbene acid was not detected in *C. cajan* seeds. *C. cajan* seeds are not the accumulating part for cajaninstilbene acid.

Taken together, the overall contents of cajaninstilbene acid in leaves were much higher than those in other parts of *C. cajan* seedlings. In a comprehensive consideration of the temporal-spatial variation in the content of cajaninstilbene acid, *C. cajan* leaves accumulated the highest amount of cajaninstilbene acid, and the optimum harvesting time was the 35th day after the emergence of leaves.

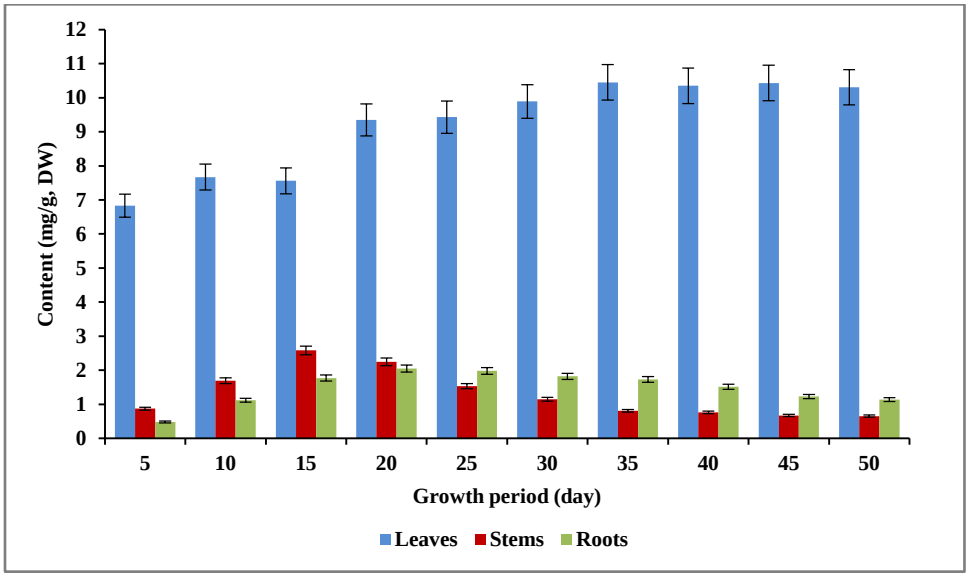


Fig. 29. Variation in content of cajaninstilbene acid in seedling leaves, stems and roots of *C. cajan* during the growth period. Each value is the mean of three replicates ± SD.

2.3 Elicitation studies in cell suspension cultures of *C. cajan*

Cell suspension culture is a very good plant material to be used for breeding, studying secondary metabolite biosynthesis and secondary metabolite production. However, cajaninstilbene acid has not been detected in cell suspension or callus cultures so far. Some of the strategies used to stimulate cajaninstilbene acid production in cell cultures involved medium modifications and a variety of explants. In our previous studies, a stabilized cell suspension culture of *C. cajan* was well established and elicitation has been employed for inducing and/or improving secondary metabolite production in the cell cultures. It would be interesting to observe elicitation effects on secondary metabolite production in *C. cajan* cell cultures.

In this study, since the stimulatory effect is dependent on the amount of elicitor added, all the elicitors (chitosan, methyl jasmonate, yeast extract and sodium chloride) were screened at various concentrations for their influence on cajaninstilbene acid production in *C. cajan* cell suspension cultures. Seven days after inoculation the suspensions which corresponded to early linear growth phase were incubated in the presence of elicitors for different periods of time. Three individual cultures were used for each experiment. Control cultures were treated with distilled water (2 ml) or ethanol (400 µl). The details of each test are shown in Table 8.

For cajaninstilbene acid identification, 80% EtOH extracts were investigated. Characteristic signals for cajaninstilbene acid in HPLC were absent from both control and elicitor-treated cell cultures. All elicitors used showed no significant increase in metabolites production. Although, environmental stress or elicitation appear to be commonly a direct stimulus for enhanced secondary metabolite production in plants or cell cultures it appeared that in *C. cajan* cell suspension cultures the biotic stress did not have any activating or stimulating effect on cajaninstilbene acid production.

The response to elicitation is dependent on the growth phase of the culture. Cells at different stages of growth have different levels of mRNA and proteins, and elicitor treatment on different growth phases may yield varied response in terms of cell growth and secondary metabolite production (Chong et al., 2005). The selection of different points of culture age was based on the growth curve of the cell suspension cultures, thus 7, 12 and 18 days old cultures are the early, middle and late, respectively, growth phases, while 22 days old cultures represent the early stationary phase. Several workers studied the optimum age of the culture for elicitation in different plant cell systems. The young culture age was more sensitive to the effect of elicitation of saponin production in *Panax ginseng* treated with yeast extract and methyl jasmonate (Lu et al., 2001). Ketchum et al. (1998) suggested a greater effectiveness of the elicitor on *Taxus* cell cultures when the elicitor was added 7 or 8 days after inoculation. The culture required about one week to recover from the stress of being transferred to fresh medium.

Stimulation of the biosynthesis of constitutive secondary metabolites during the exponential or stationary stages of cellular growth from cell tissues or upon induction by elicitation has been reported. The accumulation of the constitutive triterpene acids ursolic and oleanolic acids in *Uncaria tomentosa* cell cultures increased by elicitation during the stationary stage (Flores-Sanchez et al., 2002), while in *Rubus idaeus* cell cultures

increasing amounts of raspberry ketone (p-hydroxyphenyl-2-butanone) and benzalacetone were observed during the exponential stage. Also, induction by elicitation of secondary metabolite biosynthesis, such as the stilbene resveratrol in *Arachis hypogaea* (Rolfs et al., 1981) and *Vitis vinifera* (Liswidowati et al., 1991) cell cultures has been reported.

As cajaninstilbene acid is a constitutive secondary metabolite in *C. cajan*, a time course was made after induction with chitosan (20 mg/l), methyl jasmonate (100 µM) and yeast extract (2 g/l). All are known to induce the plant defense system. These elicitors were used to induce the metabolism of the cell cultures during the exponential and stationary phases of cellular growth. As shown in Fig. 30, the cellular growth was not significantly affected by the treatments. The results revealed that the continuous increase in dry weight, which was observed with control cell cultures, was delayed by three days in yeast extract treated cultures. The effect of elicitor treatment in form of cell growth retardation has been reported in *Salvia miltiorrhiza* cell suspension culture (Chen et al., 2000). However, no peaks for cajaninstilbene acid in HPLC were detected during the time course of the elicitation of cell cultures. Analysis by LC-MS of the extracts reveled similar results (Fig. 31).

Table 8. Elicitors applied to *C. cajan* cell cultures and harvest time.

Elicitors	Concentrations				Harvest time (days)	Cajaninstilbene acid detection
Chitosan (mg/l)	10	20	50	100	Every 2 days	Not detectable
Methyl jasmonate (µM)	50	100	200	400	Every 2 days	Not detectable
Sodium chloride (%, w/v)	0.2	0.5	1	2	Every 2 days	Not detectable
Yeast extract (g/l)	1	2	4	8	Every 2 days	Not detectable

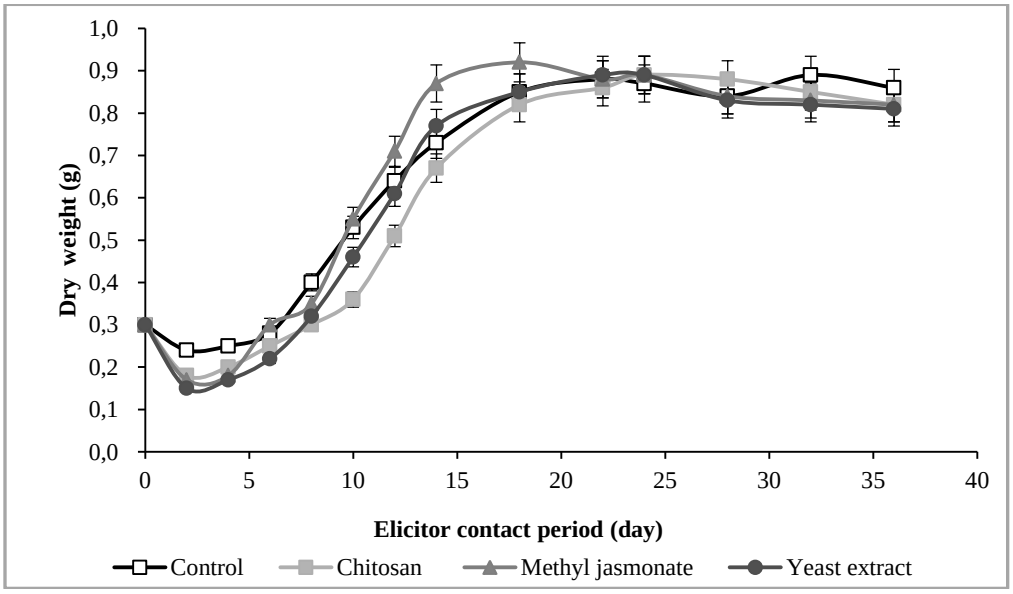


Fig. 30. Effect of elicitors on dry weight of *C. cajan* cell suspension cultures after treatment with chitosan (20 mg/l), methyl jasmonate (100 µM) and yeast extract (2g/l).

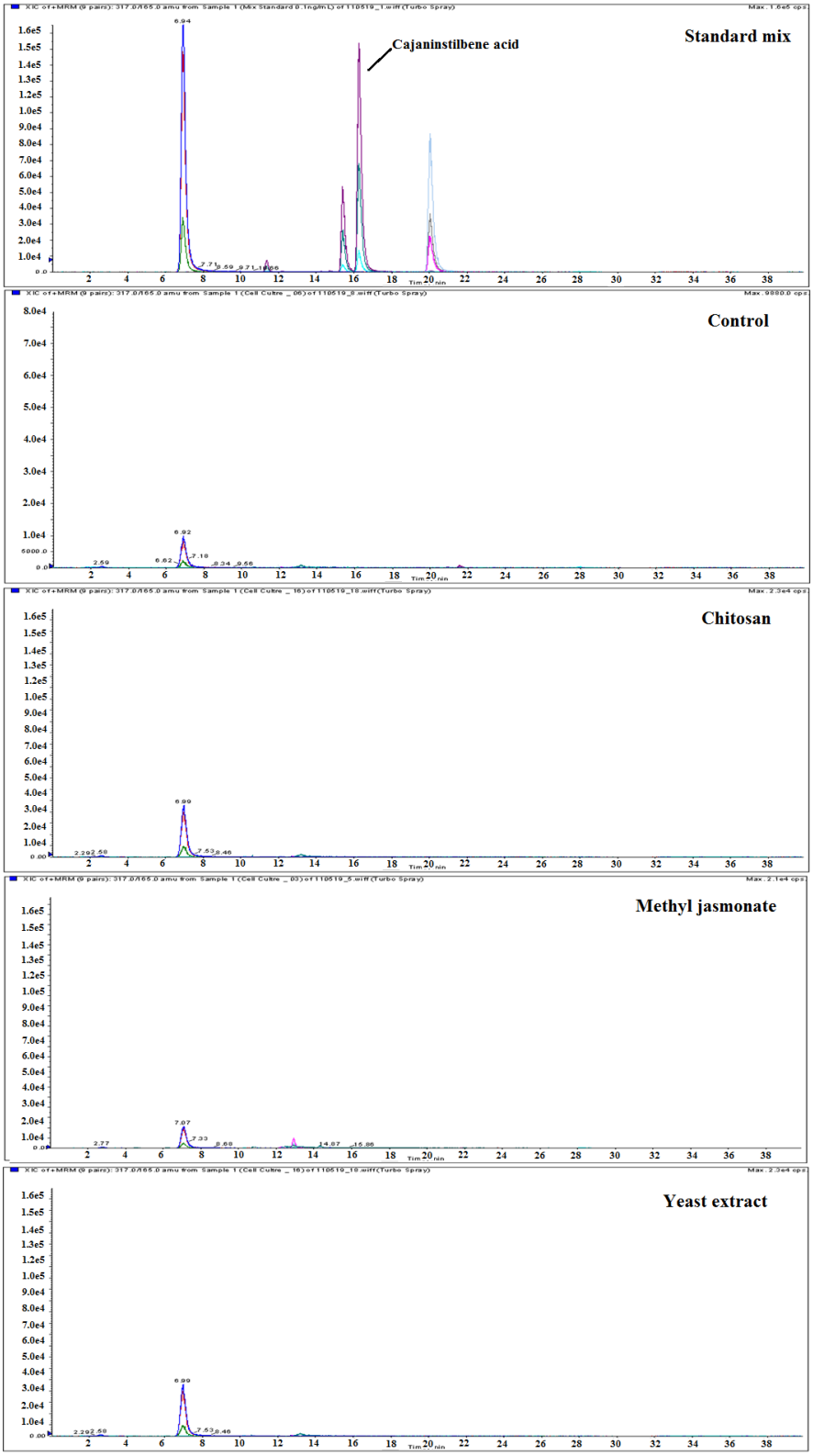


Fig. 31. LC-MS analysis of extracts from *C. cajan* cell suspension cultures after 7 days treatment with chitosan (20 mg/l), methyl jasmonate (100 µM) and yeast extract (2 g/l).

3. Detection of stilbenecarboxylate synthase (STCS) activity in crude protein extracts from *C. cajan* leaves

In previous studies, cajaninstilbene acid was detected only in *C. cajan* seedlings, with *C. cajan* leaves accumulating the highest amount of cajaninstilbene acid as the most predominant phytochemical. So, well established *in vitro* leaves were used to prepare crude protein extract for the detection of stilbenecarboxylate synthase (STCS) activity. Protein extraction from intact plant tissues is typically more challenging than from cell cultures and the extraction of high-quality protein becomes even more complicated when proteins are extracted from mature tissues, which typically contain lower concentrations of proteins and higher levels of interfering compounds. In these cases, as the protein extraction from *C. cajan* leaves has not been studied before, the optimization of protein extraction is essential to the following enzyme assays.

3.1 Efficiency of protein extraction solutions

In order to optimize protein yields from small amounts of *C. cajan* tissues, the performance of six commonly used proteins extraction buffers with various pH values (Table 9) was tested for mature leaf proteins. All extraction solutions were supplemented with high concentration of DTT to prevent phenolic oxidation of proteins and with high amount of Polyclar AT to absorb phenolic compounds. Furthermore, the extraction was done twice in order to remove phenolic compounds from the samples more efficiently by Polyclar AT. The quantitative comparison of the performance of the extraction solutions is illustrated in Table 9. SDS-PAGE evaluation of extracted proteins is shown in Fig. 32.

Table 9. The performance of different protein extraction buffers with various pH values.

Buffer	pH	Protein yields (mg /g fresh weight)
Sodium citrate	4	5.79
Sodium acetate	5	6.86
HEPES	7	8.23
Sodium phosphate	7	7.76
Tris–HCl	8	4.69
Sodium borate	9	3,22

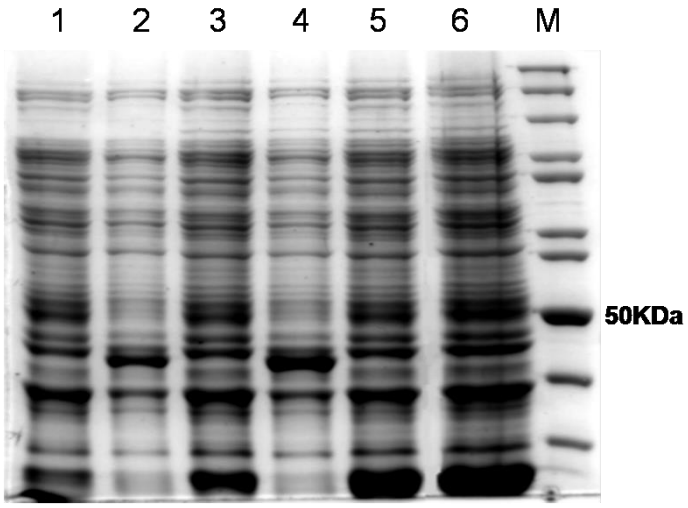


Fig. 32. SDS-PAGE evaluation of proteins extracted from *C. cajan* leaves using different buffers. 1. Sodium citrate pH 4; 2. Sodium borate pH 9; 3. Sodium acetate pH 5; 4. Tris-HCl pH 8; 5. Sodium phosphate pH 7; 6. HEPES pH 7.

As shown in Table 9 and Fig. 32, soluble proteins can be extracted by all buffers tested. However, the efficiency of protein extraction was clearly dependent on the composition of the extraction solution. Especially, the pH appeared to be a significant factor because the best protein yields were obtained with buffers in the range of pH 5-7, such as HEPES, sodium phosphate, and sodium acetate. Clearly, the highest amounts of proteins from mature *C. cajan* leaves were extracted with HEPES buffer at pH 7, with protein yields of 8.23 mg/g of fresh tissue powder.

3.2 Detection of stilbenecarboxylate synthase (STCS) activity

Incubation for 1 h of crude protein extract from *C. cajan* leaves was done with a starter CoA-ester (cinnamoyl-CoA, *p*-coumaroyl-CoA, dihydro-cinnamoyl-CoA and dihydro-*p*-coumaroyl-CoA) and malonyl-CoA as extender molecule. The procedure of enzyme assay is described in II.2.2.3. Enzyme assays containing heat-denatured crude proteins were carried out as control assays. Analysis of the enzymatic products was performed by HPLC. In addition to STCS activity, some related plant type III PKS activities, such as CHS and STS activities, might also be found in the enzyme assays according to the substrates and the proposed polyketide pathway. The products that might appear in the enzyme assays are summarized in Fig. 33.

All the enzyme assays failed to form the expected enzymatic products. A small peak (R_t = 13 min) was found in HPLC but was unrelated to the expected products after further identification. Characteristic product peaks in HPLC were absent from all enzyme assays with crude protein extract, both control and normal test. Higher protein concentrations (up to 100 µg), different reaction times (10, 20, 30, 120 min) and different pH values of reaction buffer were also tried in the assay, but no product was detected. As an alternative approach of product detection, the standard assay was incubated with [2-^{14}C] malonyl-CoA, which would result in the formation of ^{14}C-labeled products detectable by

radioactivity detector-coupled HPLC analysis. However, no product could be detected, as shown in Fig. 34.

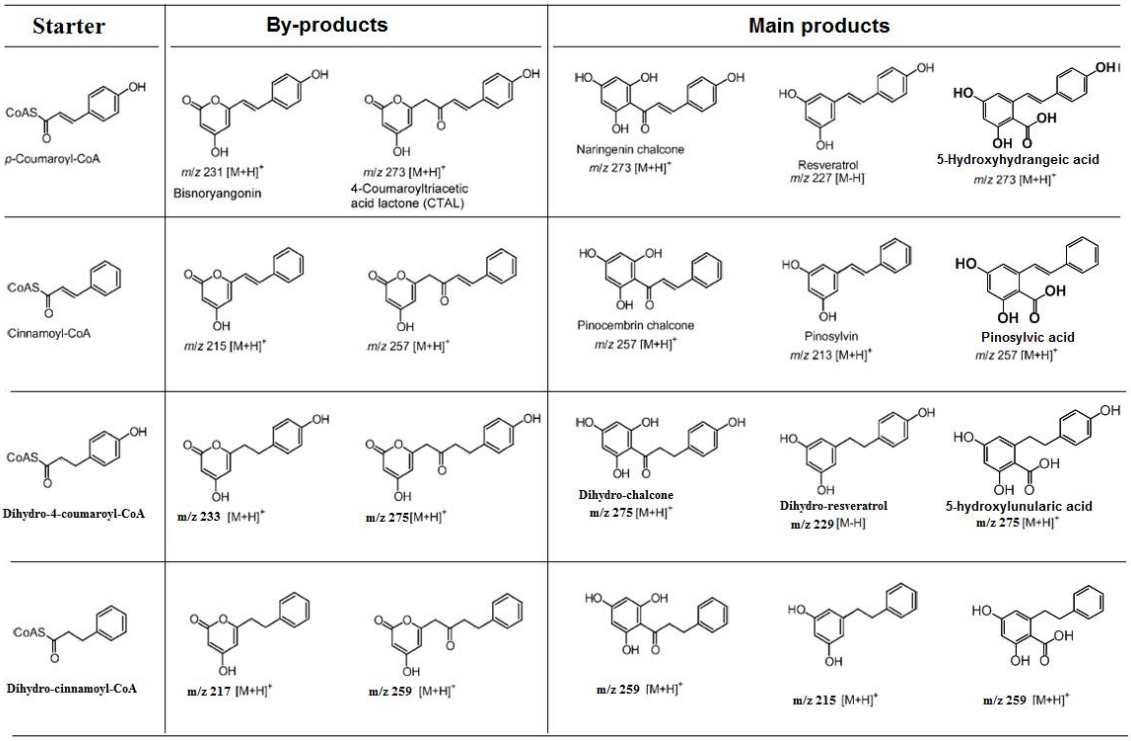


Fig. 33. Products that might be formed in the PKS assay.

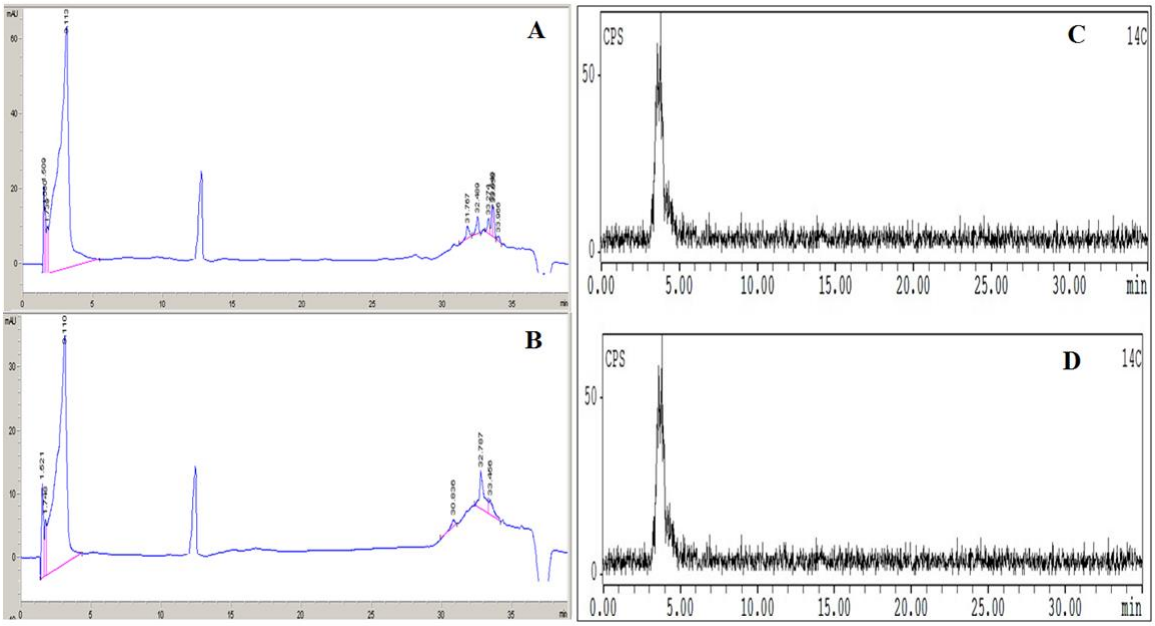


Fig. 34. Analysis of products from enzyme assays with crude protein extract using DAD-coupled HPLC (A and B) and radiodetector-coupled HPLC (C and D). A and C: control assays; B and D: standard assays

Although the results were not expected, protein extraction from intact plant tissues for enzyme assay are typically more challenging and it is hard to say what causes this problem. Failure to detect some enzyme activities on the biochemical level can be attributed to many reasons. Although extraction with optimized buffers leads to relatively high protein yields from leaves of plants, the extracts may still have too low protein contents for detection of enzyme activities and the samples need to be concentrated. In addition to having relatively low protein contents, plant tissues are often rich in proteolytic enzymes and possess high levels of non-protein contaminants that co-extract with proteins and can severely hamper protein stability, protein activity, and further analysis. These contaminants include cell wall and storage polysaccharides, organic acids, salts, pigments, phenolic compounds, and a broad array of other secondary metabolites. The extraction of high-quality protein becomes even more complicated when proteins are extracted from mature tissues, which typically contain lower concentrations of proteins and higher levels of interfering compounds. As a consequence, more efforts need to be invested into this work in the future. This will help in understanding the biosynthesis of stilbenecarboxylate and other secondary metabolites in *C. cajan.*

4. Molecular cloning of cDNAs encoding stilbenecarboxylate synthase (STCS) and other type III polyketide synthases from *C. cajan* leaves

4.1 Candidate sequences in *C. cajan* genome

In plants, several type III PKSs have been found and all of them participate in the biosynthesis of secondary metabolites. As a ubiquitous enzyme in higher plants, CHS is the most studied enzyme, which provides the first committed step in flavonoid biosynthesis. Thus, the group of PKSs is often called the CHS super family of type III PKSs. It is known that CHSs share 44–95% amino acid sequence identity to other members of type III PKSs which utilize a variety of different substrates or use the same substrate as CHS but undergo different numbers of condensation and/or different cyclization of the polyketide intermediate. Therefore, all CHS-related candidates of type III PKSs from the genome database seem to be screened with CHS as a probe. In these cases, based on the similarity of the substrate structure and the nature of the catalytic reaction, CHS2 from *Medicago sativa* was selected as a probe to screen the *C. cajan* genome (Varshney et al., 2011) database for homologs. The database is available online on a single website (http://www.icrisat.org/gt-bt/iipg/Home.html).

A total of 12 type III PKS genes were detected in the *C. cajan* genome. One of them, C.cajan_47702, is only 255 bp (85 aa) in length and was excluded from the further analysis. The results of alignment of amino acid sequences of the 11 remaining candidates with CHS2 (*Medicago sativa*), AtPKS_A (*Arabidopsis thaliana*) and STCS (*Hydrangea macrophylla*) are presented in Fig. 35. A phylogenetic tree was generated based on this alignment (Fig. 36). As shown in the tree, the candidates were subdivided into three groups. At the bottom, two *C. cajan* sequences (C.cajan_21074 and C.cajan_19240) grouped together with AtPKS_A, which has been shown to be a anther-specific type III polyketide synthase involved in the biosynthesis of sporopollenin biosynthesis during pollen development (Dobritsa et al., 2010). This group of type III PKSs has often very low sequence identity (37-45%) to other members of this superfamily (Fig. 35). In the middle

cluster of the tree, MsCHS2 and four *C. cajan* sequences (C.cajan_24072, C.cajan_23441, C.cajan_03204 and C.cajan_20347) grouped together to form a clearly separated cluster. The four *C. cajan* sequences share 76-91% identity at the amino acid level, they are most likely to be chalcone synthases. C.cajan_37423 is between the CHS-cluster and anther-specific type III PKS cluster, it has 79.4% identity to MsCHS2, but C.cajan_37423 with a length of 500 bp was a middle fragment without starting and ending sequences. At the top of the tree, another clearly separated cluster appeared including HmSTCS and three *C.cajan* sequences (C.cajan_43674, C.cajan_45286 and C.cajan_08958). The three *C.cajan* seuences showed similar identity to MsCHS2 and HmSTCS in the range of 63.8-67.9%. C.cajan_38726 is between cluster STCS and CHS, its function is hard to predict only from sequence analysis. Consequently, eight candidates of type III PKSs from *C. cajan* were selected for further functional analysis. The corresponding primer sets are mentioned under II.1.5.2. Amplification of the full lengths of these candidates was carried out in the following work.

Percent Identity

Divergence

	1	2	3	4	5	6	7	8	9	10	11	12	13	14		
1		89.7	88.1	91.5	76.4	80.5	70.4	67.9	65.3	64.0	79.4	39.3	41.3	38.6	**1**	MsCHS2
2	11.1		93.6	89.5	79.0	84.1	72.4	70.4	67.6	66.6	85.5	40.7	41.5	40.4	**2**	C.cajan_23441
3	12.9	6.7		90.2	81.8	84.6	73.2	72.2	68.6	67.9	87.3	40.2	41.5	39.8	**3**	C.cajan_03204
4	9.0	11.4	10.5		77.7	82.8	71.5	70.3	67.7	66.4	81.2	39.0	41.3	38.5	**4**	C.cajan_24072
5	28.4	24.7	20.9	26.5		73.1	69.1	66.1	66.1	63.0	73.9	38.7	41.4	38.6	**5**	C.cajan_20347
6	22.7	18.0	17.3	19.6	33.4		75.1	73.0	68.6	68.4	77.0	41.4	41.3	41.4	**6**	C.cajan_38726
7	37.6	34.4	33.2	35.9	39.8	30.4		66.1	63.8	64.5	65.5	40.7	38.8	39.3	**7**	HmSTCS
8	41.9	37.6	34.7	37.9	45.0	33.6	45.0		81.8	82.0	68.5	39.3	39.3	39.1	**8**	C.cajan_43674
9	46.4	42.3	40.5	42.2	45.0	40.6	49.2	20.9		76.6	67.9	39.0	38.5	37.2	**9**	C.cajan_08958
10	48.7	44.1	41.9	44.4	50.7	41.0	47.8	20.7	28.0		63.0	36.9	37.0	36.5	**10**	C.cajan_45286
11	24.2	16.2	14.0	21.7	32.0	27.6	46.1	40.8	41.8	50.6		44.2	42.4	43.6	**11**	C.cajan_37423
12	114.3	108.9	110.9	115.4	116.7	106.4	108.8	114.3	115.4	124.1	96.7		66.3	78.6	**12**	C.cajan_19240
13	106.5	106.1	106.1	106.5	106.3	106.5	116.1	114.3	117.3	123.8	102.7	44.6		66.8	**13**	C.cajan_21074
14	116.8	110.1	112.1	117.2	117.0	106.3	114.0	115.1	122.5	125.9	98.7	25.3	43.6		**14**	AtPKS_A
	1	2	3	4	5	6	7	8	9	10	11	12	13	14		

Fig. 35. Sequence identity of type III PKS candidates from the *C. cajan* genome with CHS (*Medicago sativa*), AtPKS_A (*Arabidopsis thaliana*) and STCS (*Hydrangea macrophylla*)

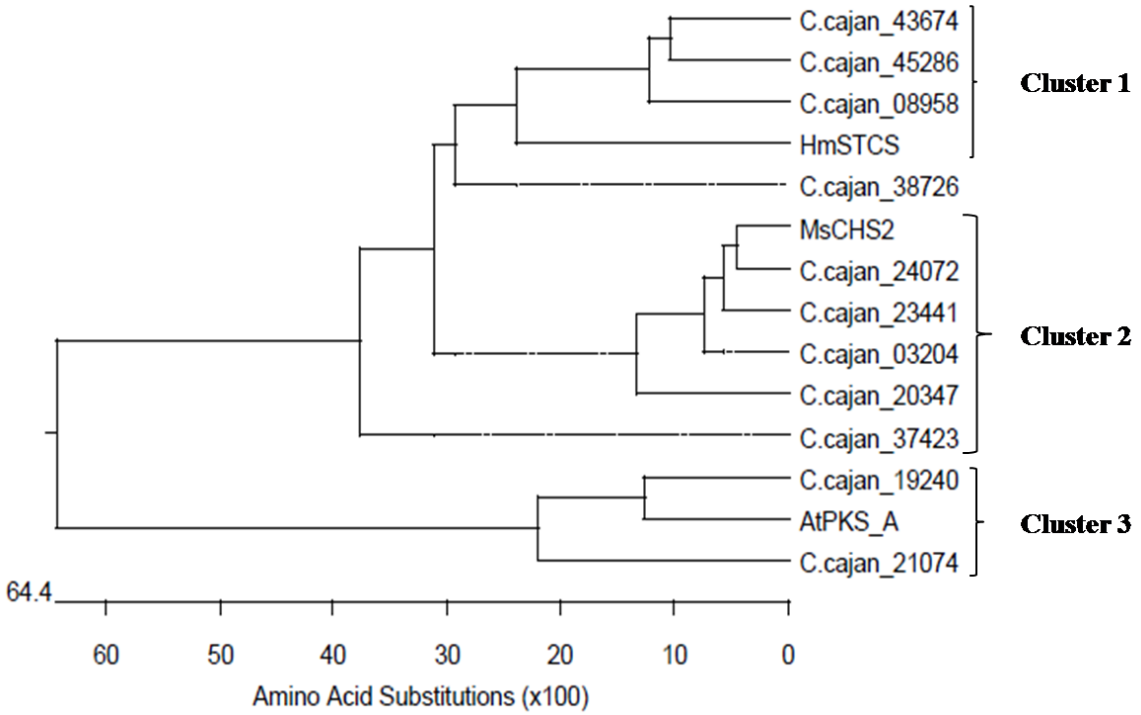


Fig. 36. Phylogenetic tree analysis of type III PKS candidates from *C. cajan*. A total of 11 type III PKS sequences of *C. cajan* were aligned by the clustal W method. The tree was constructed using the CLUSTAL W 1.8 program. CHS (*Medicago sativa*), AtPKS_A (*Arabidopsis thaliana*) and STCS (*Hydrangea macrophylla*) were employed as references.

4.2 Amplification of cDNAs encoding PKSs

A cDNA pool was prepared by reverse transcription of RNA isolated from 25-day-old *C. cajan* leaves and used as a template for PCR reaction (II.2.3.3). The obtained cDNA was examined for its quality by checking the successful expressionof 18s rRNA. Full length genes were amplified using gene specific primers based on the obtained open-reading frame (ORF) sequences. The primer sets used for amplification of the candidates are mentioned under II.1.5.2. The cDNAs were amplified using proof-reading Phusion DNA polymerase Hot start II and the PCR program listed under II.2.3.5.
In case of candidates from group one of the phylogenetic tree, three candidates were renamed pinosylvic acid synthases, PAS1, PAS2 and PAS3 (Table 10). The forward and reverse gene-specific primer pairs given in II.1.5.2 were used. In all three candidate sequences, restriction sites for *Nhe* I *and Xho* I were absent. So, the restriction site of *Nhe* I was integrated into the forward primers, while the restriction site of *Xho* I was integrated into the reverse primers. With the standard PCR programme and an annealing temperature of 58°C, two PCR products of 1182 and 1188 bp for PAS1 and PAS3, respectively, were obtained (Fig.37). Amplification of PAS2 failed, despite variation in the PCR conditions. No product was obtained, indicating that this gene was not expressed in leaves.

In case of candidates from group two of the phylogenetic tree, five candidates were

renamed C.CHS1, C.CHS2, C.CHS3, C.CHS4 and C.CHS5 (Table 10). The primers were designed in such a way that the forward primers contained a *Nhe* I restriction site, whereas the reverse primers contained an *Eco R*I restriction site. The designed primers were used to amplify the five members of group two sequences in a standard PCR at an annealing temperature of 60°C, yielding five products of 1167, 1158, 1167, 1170 and 1176 bp in length for C.CHS1, C.CHS2, C.CHS3, C.CHS4 and C.CHS5, respectively (Fig 37). All the seven amplified sequences were sequenced on both strands and used in the following work. Their nucleotide sequences are presented in the appendix.

Table 10. Renamed candidate sequences.

Candidates in the genome	Renamed as
C. cajan_45286	PAS1
C. cajan_43674	PAS2
C. cajan_08958	PAS3
C. cajan_ 03204	C.CHS1
C. cajan_ 20347	C.CHS2
C. cajan_ 23441	C.CHS3
C. cajan_ 24072	C.CHS4
C. cajan_ 38726	C.CHS5

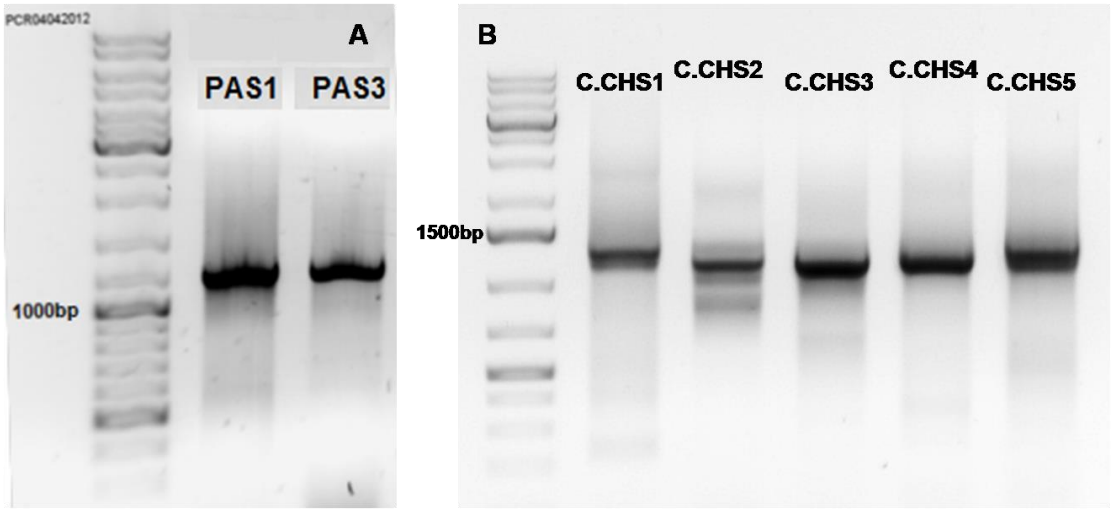


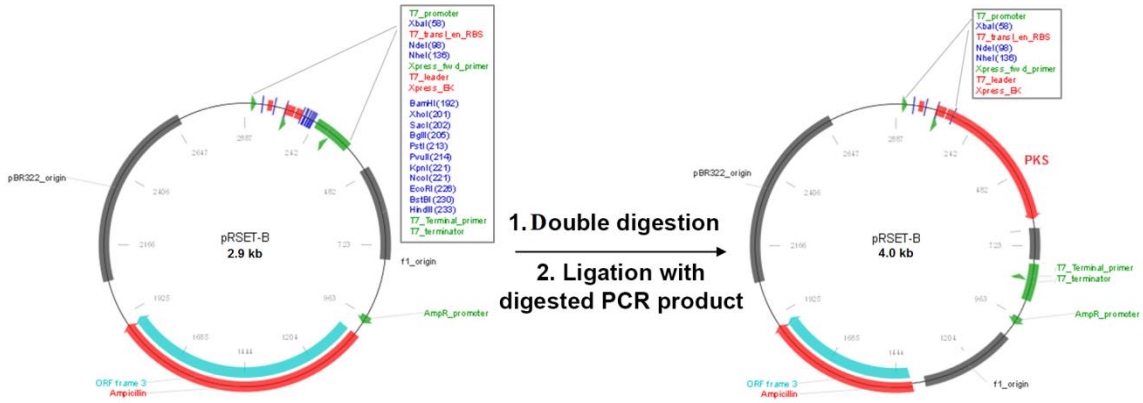
Fig. 37. Agarose gel electrophoresis of PCR products of PKS ORFs from *C. cajan*

4.3 Heterologous expression of *C. cajan* PKS ORFs in *E. coli* and protein purification

The PCR products of the aforementioned reactions were purified, digested and ligated to pRSET B expression vector between the corresponding restriction sites (II.2.3.8, II.2.3.9) (Fig. 38). The ligation products were introduced in *E. coli* DH5α (II.2.3.10) to obtain large amounts of this plasmid, followed by plasmid isolation (II.2.3.11). Confirmation of successful ligation was carried out by restriction endonuclease digestion analysis. The constructed plasmids were sequenced to ensure the presence of the right insert. No frame shifts were detected. Plasmids were transferred into *E. coli* strain BL21 (DE3) pLysS for carrying out heteologuos expression (II.2.3.12). They were functionally expressed in *E.*

coli as 6xHis-tagged proteins. Following protein extraction and purification by affinity chromatography, the efficiency of overexpression and confirmation of their production as soluble proteins were examined by SDS-PAGE (Fig. 39 and Fig. 40). Band sizes of 42-44 kDa were expected (39-41 kDa+3 kDa N-terminal tag). The Ni-NTA system was used for the purification of the 6xHis-tagged PKSs. The PKSs can be induced upon addition of IPTG and constituted major bands in the crude soluble protein. Finally, they were eluted from the Ni-NTA gel bed. The yield of pure protein was determined by the Bradford method and a certain amount of purified protein was used in the following assays of enzymatic activity.

A



B

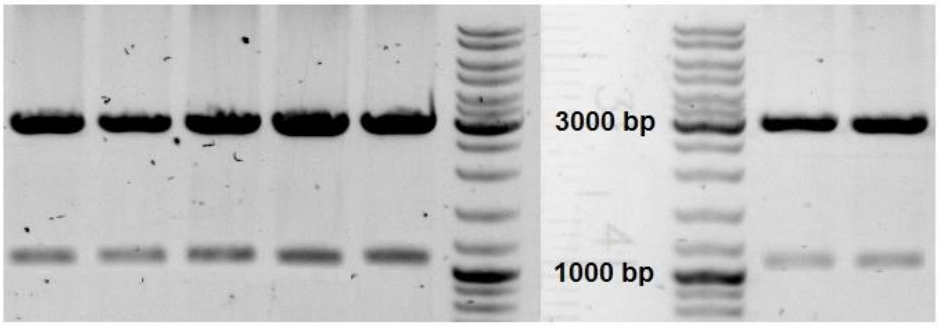


Fig. 38. Construction of the expression vector for PKS cDNAs. A. Vector maps, B. Confirmation of successful ligation and presence of inserts.

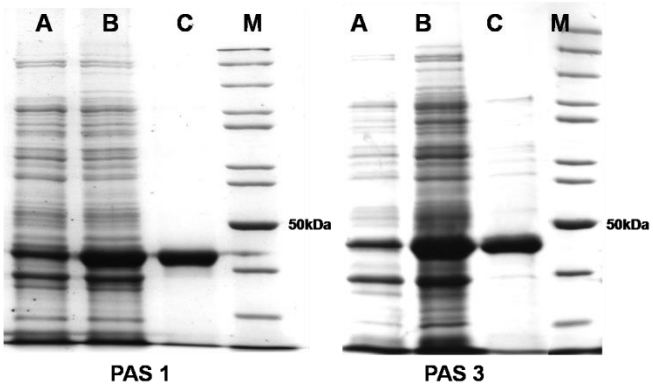


Fig. 39. SDS-PAGE of the over-expressed proteins PAS1 and PAS3.
A: Crude protein before induction with IPTG (pellets)
B: Crude protein after induction with IPTG (pellets)
C: Purified protein (≈ 5µg)
M: Marker

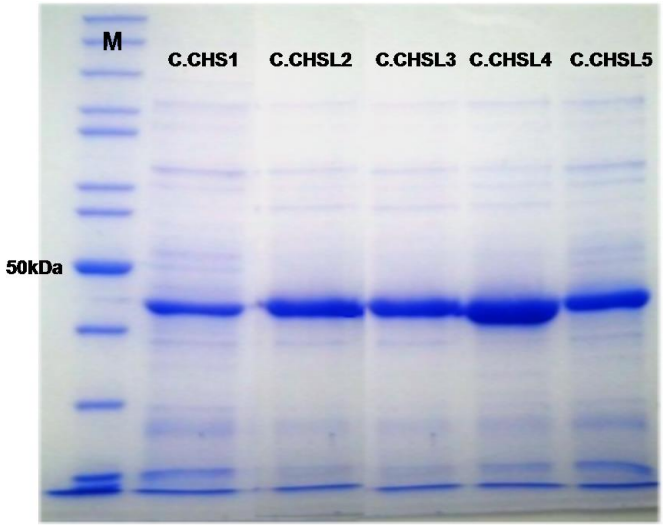


Fig. 40. SDS-PAGE of the purified proteins C.CHS1,C.CHS2, C.CHS3, C.CHS4 and C.CHS5 (≈ 5µg each).

4.4 Test for enzymatic activity

The functions of the purified fusion proteins were investigated using an *in vitro* enzymatic incubation. The procedure of enzymatic assay is described in II.2.2.3. Purified recombinant protein (10 µg) was used for PKS assays. According to the natural products identified from the plantand the proposed polyketide biosynthesis pathway, the first assays were carried out with cinnamoyl-CoA and malonyl-CoA as the starter substrates to see whether the CHS product pinocembrin or pinosylvic acid (PA) was synthesized. The stilbenecarboxylate was expected from a STCS activity assay. Confirmation of the enzymatic nature of the products was carried out doing parallel incubations containing heat-denaturated protein. After termination of incubation by acidification, the reaction products were extracted by ethyl acetate and the extracts then analyzed by HPLC. The

identity of the products was confirmed by comparing the retention time, the UV spectrum and the MS–MS spectrum with those of authentic reference materials.
Detection of pinocembrin, which is formed by nonenzymatic isomerization of pinocembrin chalcone, together with two unidentified compounds A and B as C.CHS1-5 products proved that these enzymes are functional CHSs (Fig. 41). (The R_t value of pinosylvin which is the product of STS should be 20.03 min). However, PASs gave neither pinocembrin nor pinosylvic acid and produced only compounds A and B in the enzyme assay (Fig. 42). Incubation with heat-denatured protein showed no peaks in the HPLC chromatogram, indicating the enzymatic origin of the products.

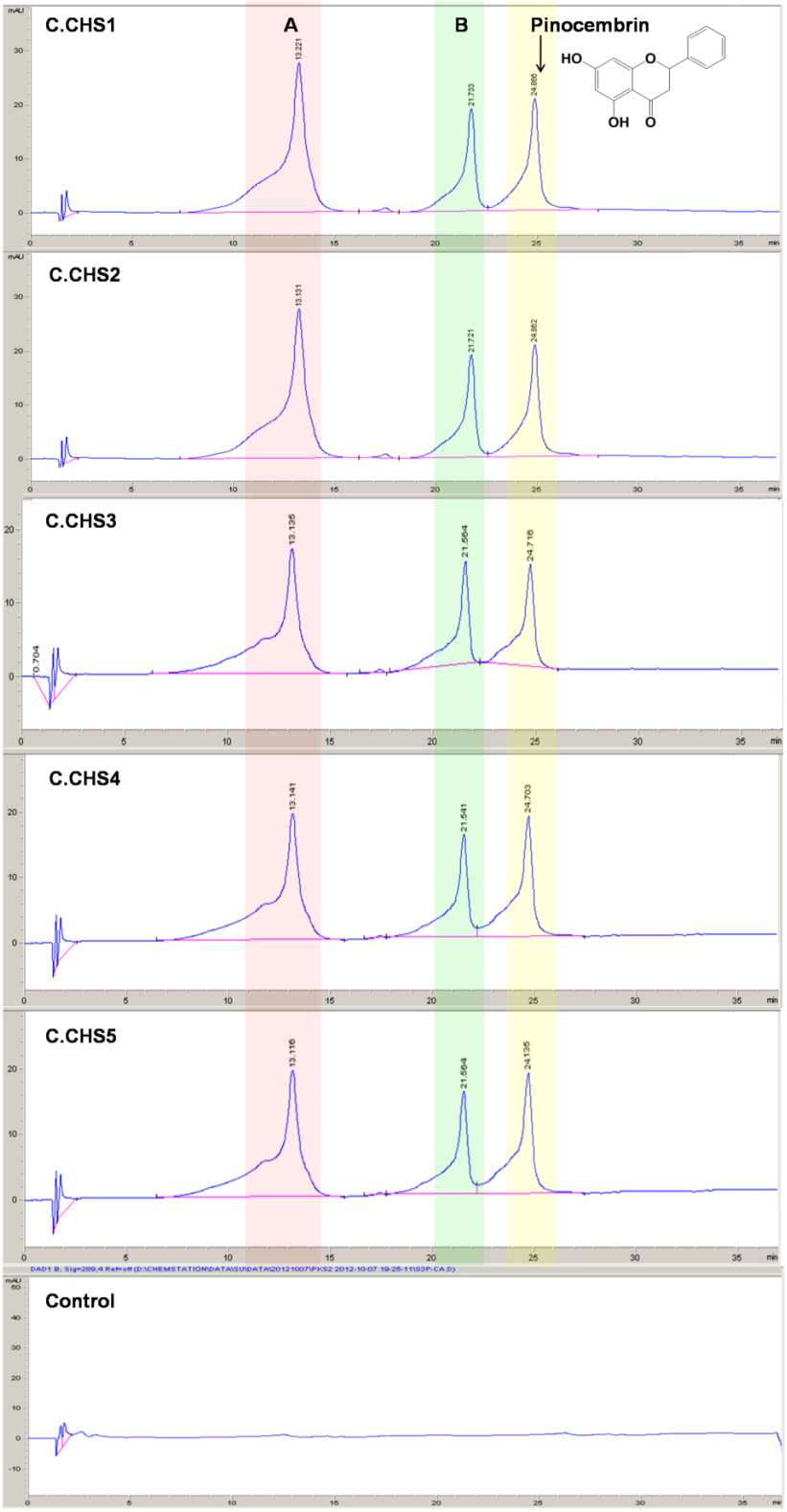


Fig. 41. HPLC analysis of C.CHSs assays containing cinnamoyl-CoA and malonyl-CoA. Incubation with denatured protein served as control. The UV detection wavelength was 280 nm. A and B, products identified below.

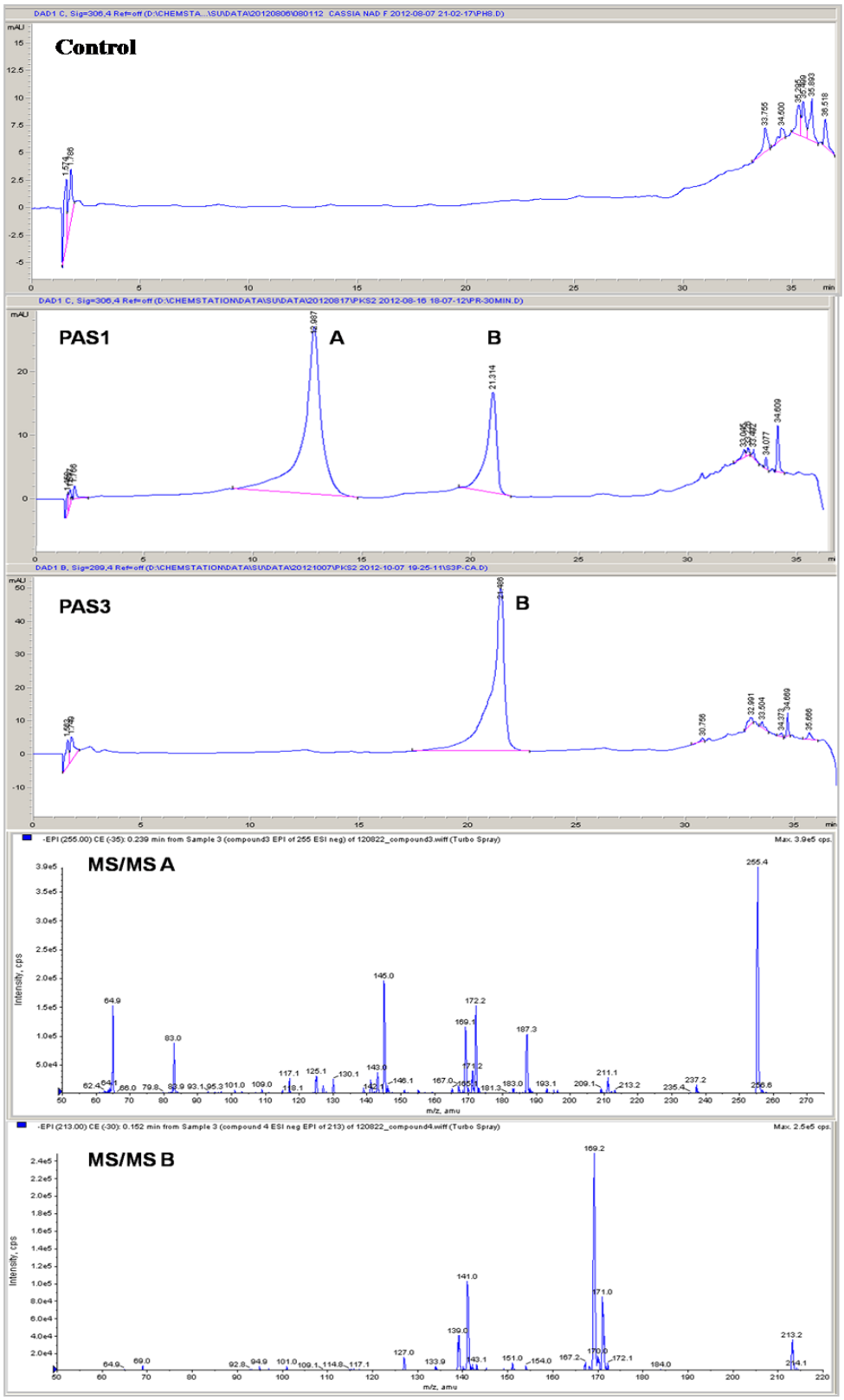


Fig. 42. HPLC analysis of PAS assays containing cinnamoyl-CoA and malonyl-CoA and LC-MS/MS spectra of compounds A and B.

In order to elucidate the structures of compounds A and B, large-scale incubations were conducted. Compounds A and B were separated and purified by preparative HPLC and analyzed by LC-MS/MS. Spectral data revealed that compound B was styrylpyrone (Fig. 42), a triketide formed from cinnamoyl-CoA and only two molecules of malonyl-CoA. It has the same reaction pathway as bisnoryangonin, which has been reported to be a CHS by-product *in vitro* (Akiyama et al., 1999), but arising from a different starter. Multi-stage LC-MS (negative-ion mode) gave a $[M\text{-}H]^-$ peak at m/z 213 in MS, $[M\text{-}H\text{-}CO_2]^-$ at m/z 169 in MS/MS (precursor ion at m/z 213) and $[M\text{-}H\text{-}CO_2\text{-}CH_2CO]^-$ at m/z 127, showing successive cleavage of the α-pyrone ring. Further identification of compound B as styrylpyrone was derived from the retention time and the UV–vis spectrumin comparison to the data of the reference.

LC-MS of compound A gave a parent peak at m/z 255, suggesting that its molecular mass was 256 Da, i.e. 42 mass units (CH_2-CO) larger than that of styrylpyrone. This would be expected for both pinosylvic acid synthesized by a STCS reaction and cinnamoyl triacetic acid lactone (CITAL, tetraketide), the pyrone product after three condensations without CHS- or STS- type ring folding (Fig. 42). Multistage LC-MS gave $[M\text{-}H\text{-}CO^2]^-$ at m/z 211 in MS/MS and $[M\text{-}H\text{-}CO_2\text{-}CH_2CO]^-$ at m/z 169, which is the same successive fragmentation patternas for styrylpyrone, indicating that compound A also had an α-pyrone ring. It also showed a base peak ion at m/z 130 that was typical for the cinnamoyl moiety ($[C_6H_4CH{=}CH{-}C{=}O]^-$). Comparison of these spectral data with the structures showed that they could originate only from CITAL and not from pinosylvic acid, indicating that the enzyme did not synthesize a stilbenecarboxylate from cinnamoyl-CoA.

Taken together, these independent findings from different approaches suggested that *C. cajan* expressed two polyketide synthases of the CHS-like protein family: CHSs as expected for flavonoid biosynthesis, and PASs that performed two or three condensations and released the corresponding lactones (Fig. 43). Similar results were reported in Akiyama's studies (Akiyama et al., 1999). In their opinion, although CITAL and styrylpyrone are common by-products of all CHSs and STSs, their productions are limited to *in vitro* reactions. In our case, there are at least three good reasons to believe that PASs are distinct enzymes which catalyze two or three consecutive decarboxylative condensations of cinnamoyl-CoA and malonyl-CoAs without aromatic ring formation and PASs gave neither CHS nor STS activity in the enzyme assay. The first is that leaves of *C. cajan*, from which the cDNA was cloned, contain stilbenecarboxylate derivatives such as cajaninstilbene acid. The second is that this cDNA was cloned together with C.CHS, the normal CHS, and the third is that PASs protein produced CITAL and styrylpyrone as the major products *in vitro*. It is interesting to note that PAS3 produced only styrylpyrone, and this kind of enzyme has not been reported before. This is the first cDNA encoding an enzyme with such a function. With all this evidence we propose that PAS protein is a novel plant PKS that catalyzes only chain elongation without cyclization *in vitro* and may be related to the biosynthesis of stilbenecarboxlate.

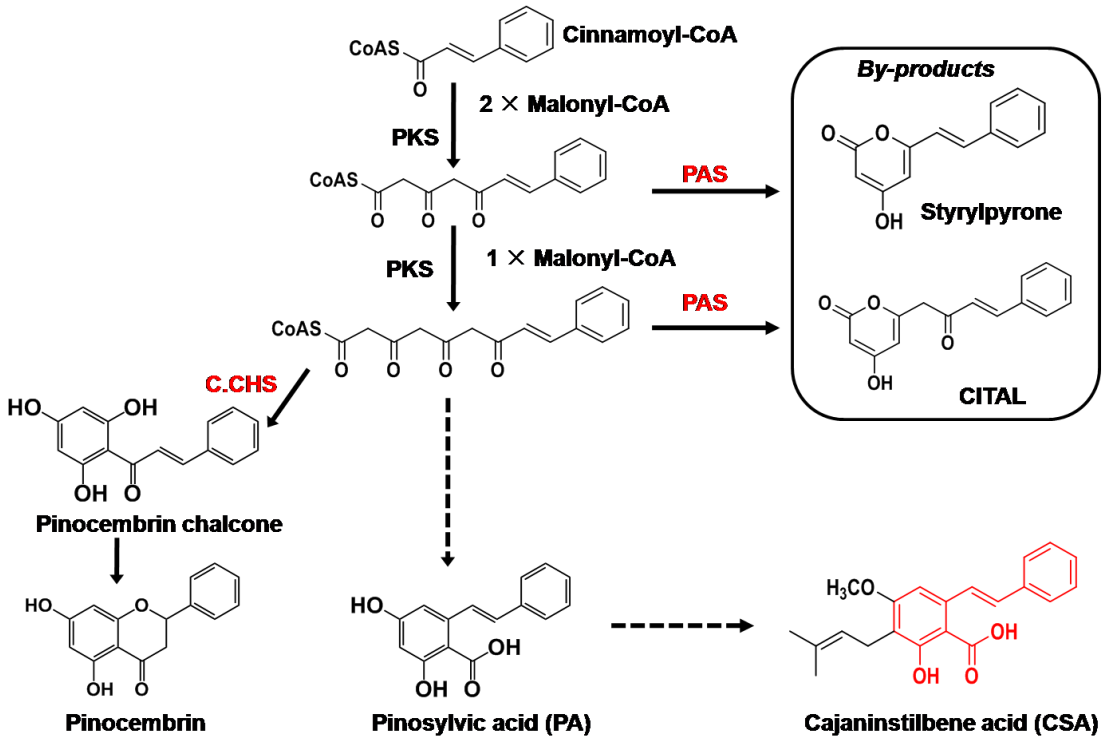


Fig. 43. Reactions catalyzed by PKSs from *C. cajan*. Solid lines indicate steps detected in this work. Dashed lines are proposed steps.

Despite considerable efforts, detection of the enzymatic activity of STCS, which is believed to be involved in stilbenecarboxylate biosynthesis, failed. Some possible explanations might be as follows. One is the presence of another PKS which is specific for stilbenecarboxylate biosynthesis and could notbe obtained by PCR under the conditions employed in this study. The second possibility is that the presence of a specific cyclase, which coacts with PAS to produce stilbenecarboxylate, could account for the biosynthesis of pinosylvic acid *in vivo*, similar to the situation of cannabinoid biosynthesis in *Cannabis sativa* (Gagne et al., 2012). In this case, the correct cyclization of the linear polyketide intermediate might need the help of a cyclase and PAS is assumed to produce stilbenecarboxylate only by close association with this cyclase protein. Therefore, it seemed possible that a STCS activity of the CHS-like protein was not detected because it required a polyketide cyclase, or because the *in vitro* incubations lacked some condition that was necessary for the correct function of the protein. Further investigations to uncover this mechanism will be descirbed in section III-5.

4.5 Biochemical characterization of recombinant PKSs

In the following sections, characters were determined using cinnamoyl-CoA and malonyl-CoA as substrates for PASs and C.CHSs, unless otherwise mentioned. Enzyme assays were carried out using the affinity-purified enzymes without cleaving the His-tag.

4.5.1 Biochemical characterization of PASs

4.5.1.1 Determination of pH optima

The pH of a solution can have several effects on the structure and activity of enzymes. For example, the pH can affect the state of ionization of acidic or basic amino acids. If this occurs, the ionic bonds that help to determine the 3-D shape of a protein can be altered, which can lead to altered protein properties or even an inactive enzyme. Changes in pH may not only affect the shape of an enzyme but also the shape or charge properties of the substrate so that it either cannot bind to the active site or cannot undergo catalysis. The dependence of PAS1 and PAS3 activity on pH was studied (Fig .44).
PAS1 and PAS3 had similar properties when compared to each other. The maximum catalytic activity and stability of PAS1 was at pH 6-7. PAS1 is also relatively active at lower pH conditions, about 60% activity was even observed at pH 3. The catalytic activity sharply decreased when the pH value was higher than 7. Subsequent enzyme assays were carried out at pH 6. For PAS3, the best conditions were pH 7. It also had some performance at lower pH values but different from PAS1. At the optimum pH, another series of incubations were performed at different temperatures to determine the optimum temperature.

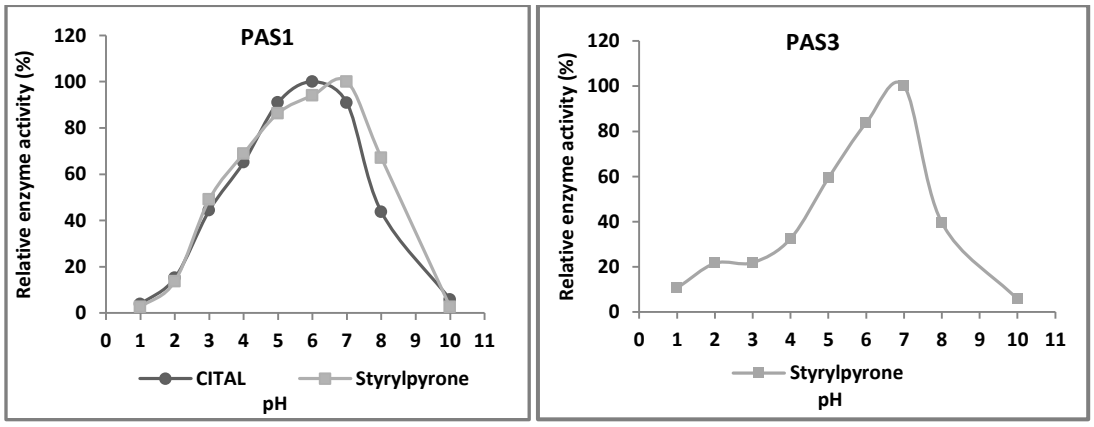


Fig. 44. Effect of pH on PAS1 and PAS3 activities.

4.5.1.2 Determination of temperature optima

All chemical reactions respond to temperature. For the same reason temperature affects enzyme reactions. Enzymes are not active at temperatures below 0°C. With the increase in temperature, the velocity of the reaction increases because more energy is supplied to stimulate molecular interactions. In case of PAS1 and PAS3, the velocity of the reaction increased with temperature until 40°C. Each enzyme has a temperature range in which a maximal rate of reaction is achieved. This maximum is called the temperature optimum of the enzyme.
The temperature optimum of both PAS1 and PAS3 was between 30-35°C. At incubation temperatures above 40°C, the proteins began to denature, lost their functions and became inactive to operate properly (Fig. 45).

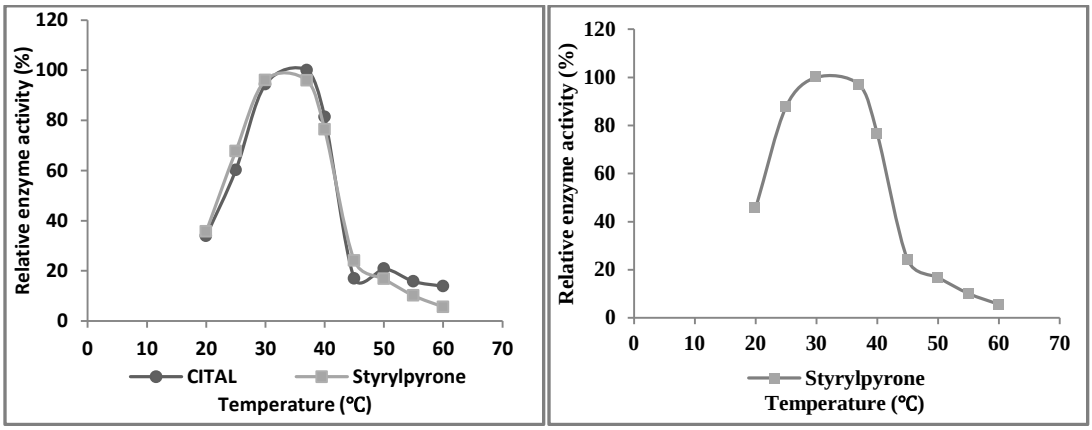


Fig. 45. Effect of temperature on PAS1 and PAS3 activities.

4.5.1.3 Linearity with protein amount

In order to study the effect of increasing enzyme concentrations on the reaction rate, care should be given to the substrate concentration not to be limiting. Any change in the amount of product formed over a specified period of time will be dependent on the level of enzyme present.

The amount of the enzymatically formed two products (nmol) was determined as a function of the protein amount in the standard assay. Products fromation of PAS1 and PAS3 was linear with the protein concentration up to 12 and 10 µg, respectively, in the standard assay (250 µl). Enzyme activities were inhibited at protein concentrations higher than 20 µg/ incubation (Fig. 46).

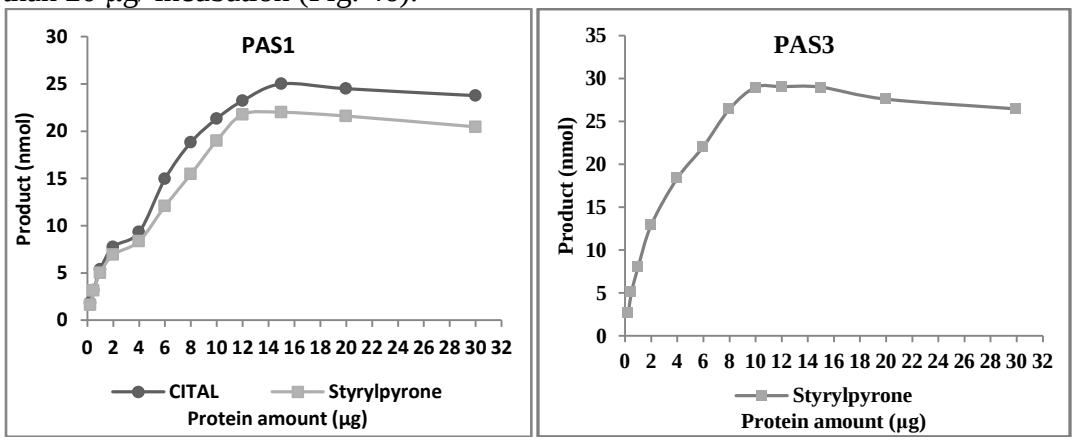


Fig. 46. Effect of protein amount on PAS1 and PAS3 activities.

4.5.1.4 Linearity with incubation time

A graph of product amount in concert with time includes three phases. At early time points, the rate of product accumulation increases linearly with time. For an extended period of time, the product amount reaches a plateau and does not change with time. At later times, the substrate is depleted, so the curve starts to level off.

Enzymatically formed products (nmol) were determined as a function of incubation time

in the standard assay. Accumulation of products from PAS1 and PAS3 was linear with incubation time up to 15 min (Fig. 47).

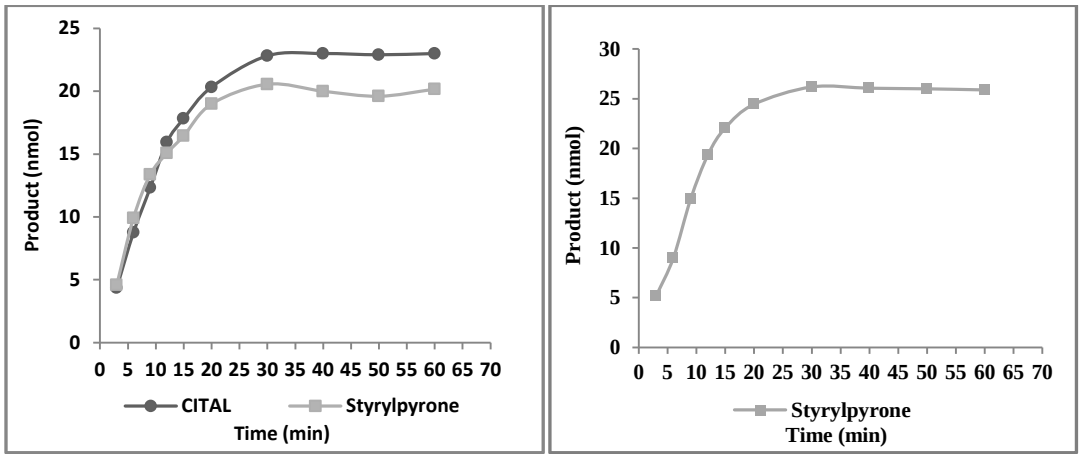


Fig. 47. Effect of incubation time on PAS1 and PAS3 activities.

4.5.1.5 Study of substrate specificity

Both PAS1 and PAS3 were tested under the respective optimum assay conditions with a series of potential substrates. The PASs were incubated in the presence of malonyl-CoA with cinnamoyl-CoA, *p*-coumaroyl-CoA, dihydro-*p*-coumaroyl-CoA, dihydro-cinnamoyl-CoA, acetyl-CoA, benzoyl-CoA, *m*-coumaroyl-CoA, *o*-coumaroyl-CoA and salicoyl-CoA. Both recombinant enzymes revealed a broad degree of substrate specificity (Table 11). The enzymes were active with most substrates; however, the enzymatic products and catalytic activities were different. An STS-type product retaining the terminal carboxyl group (STCS reaction) was not detected from any of the substrates. PASs exhibited neither CHS nor STS activities with all potential substrates and produced only the corresponding triketide and tetraketide lactones in the enzyme assay. One possible reason for its relaxed substrate specificity is that PASs are distinct enzymes which catalyze only chain elongation of substrates without formation of an aromatic ring in a polyketide pathway and most substrates can be accepted in this reaction. It was suggested that the CHS-related protein provided such precursor, and the lactones were proposed to be a derailment product in absence of the subsequent reactions (Akiyama et al., 1999).

C. cajan contains a large amount of stilbenecarboxylates and their derivatives (e.g. cajaninstilbene acid), which are most likely derived from cinnamoyl-CoA according to their structure characteristics (Fig. 48). Experiments with this substrate showed incorporation rates of labeled malonyl-CoA into radioactive products that were 30-45% higher than with *p*-coumaroyl-CoA, suggesting that cinnamoyl-CoA was a better substrate than *p*-coumaroyl-CoA. However, it is noteworthy that much higher activities (compared to cinnamoyl-CoA) were observed with dihydro-cinnamoyl-CoA, dihydro-*p*-coumaroyl-CoA and *m*-coumaroyl-CoA. Especially for dihydro-*p*-coumaroyl-CoA and dihydro-cinnamoyl-CoA, Schröder's group observed a similar non-CHS type III PKS in *Hydrangea macrophylla*, which had stilbenecarboxylate synthase (STCS) activity with only dihydro-*p*-coumaroyl-CoA and produced

stilbenecarboxylate**,** in addition to triketide and tetraketide lactones. Here, the PASs accepted the substrates dihydro-cinnamoyl-CoA and dihydro-*p*-coumaroyl-CoA with relatively high catalytic activity but produced only triketide. None of the identified natural products or derivatives derived from them were known from *C. cajan*. This was not unusual or unexpected from in vitro reactions because the corresponding byproducts or derailment products should also be closely related to the natural products in plant. Low activities were found with acetyl-CoA, benzoyl-CoA, and salicoyl-CoA. No activity was detected with *o*-coumaroyl-CoA. This is of particular interest for cinnamoyl-CoA because it should be the starter for the biosynthesis of pinosylvic acid, the main stilbenecarboxylate in the proposed pathway (Fig. 48). The identities of all the reaction products were verified by HPLC analysis in comparison with authentic reference compounds and by LC-MS analysis. Product amounts were quantified by radio activity detector-coupled HPLC analysis, the HPLC chromatograms for the different starter substrates being shown in the appendix.

Table 11. Substrate specificities of PAS1 and PAS3.

Starter substrate	**Product (% of maximum each)***	
	PAS1	**PAS3**
cinnamoyl-CoA	100	100
p-coumaroyl-CoA	73.9	55.8
dihydro-cinnamoyl-CoA	281.8	332.5
dihydro-*p*-coumaroyl-CoA	171.8	178.1
m-coumaroyl-CoA	220.1	245.4
o-coumaroyl-CoA	0	0
benzoyl-CoA	36.3	54.5
acetyl-CoA	46.2	54.9
salicoyl-CoA	72.4	79.3

***Calculated from the sum of the radioactive products**

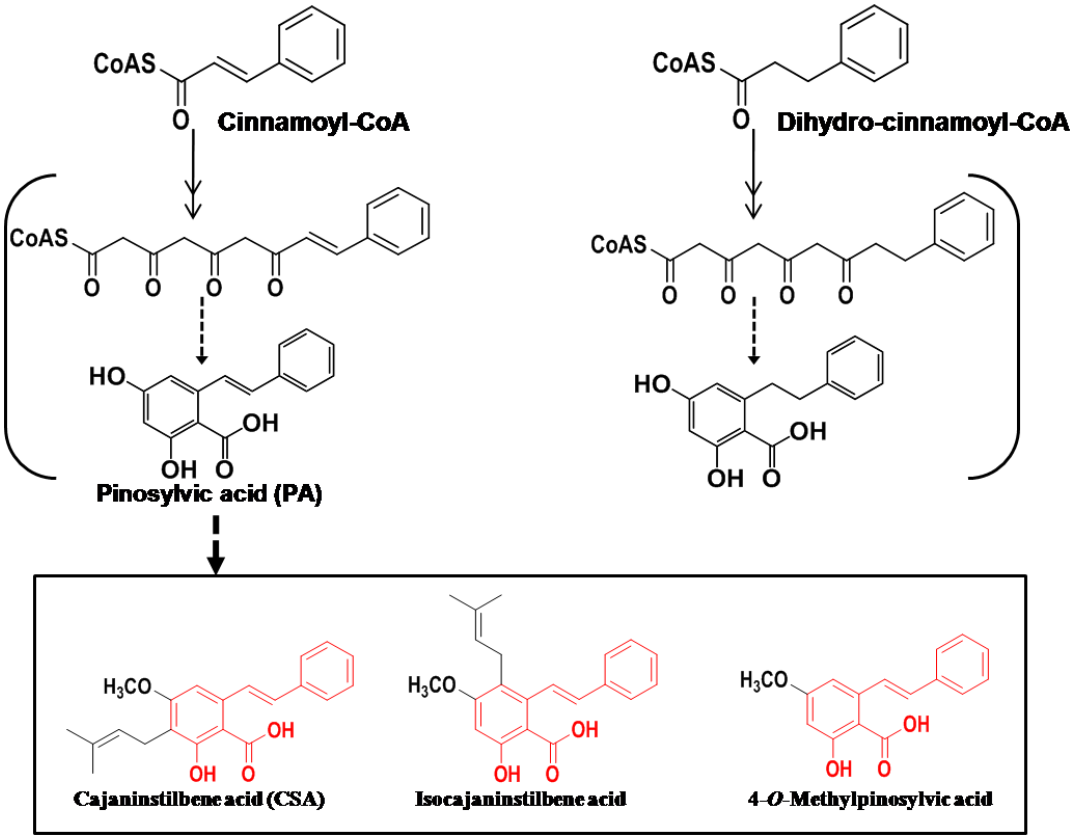


Fig. 48. Stilbenecarboxylates and derivatives in *C. cajan* and the proposed biosynthesis. The box shows the natural products identified from the plant.

4.5.1.6 Determination of kinetic parameters

The kinetic properties of PAS1 and PAS3 were determined under the optimum conditions. The assays contained 2 µg protein and the incubation period was limited to 10 min. This kinetic characterization can help in finding out the preferred substrate for each enzyme. The increase in reaction rate with increasing concentrations of substrates (cinnamoyl-CoA or *p*-coumaroyl-CoA and malonyl-CoA) obeyed Michealis-Menten kinetics (Fig. 49, 50). The Michaelis constant (K_m) allows for measuring the affinity of an enzyme for its substrate.The Lineweaver-Burk plot is a linear representation of the Michaelis-Menten equation and commonly used to determine the kinetic parameters of an enzyme. It is a plot of the reciprocal of the enzymatic reaction velocity (1/v) versus the reciprocal of the substrate concentration (1/[S]). K_{cat} is the turnover number, which reflects how efficiently an enzyme transforms a substrate once bound. The ratio K_{cat}/K_m defines a measure of the catalytic efficiency of an enzyme-substrate pair. It can be used to compare between utilization of different substrates by one enzyme and to compare the catalytic efficiency between different enzymes. Moreover, it is a good measure of substrate specificity because it assesses the interaction between the enzyme and the substrate in the ground state as well as in the transition state (Copeland, 2002). Based on this information, we can determine the best substrate for each enzyme. The summaries of the kinetic data for the two enzymes are presented in Table 12.

Table 12. Kinetic parameters of PAS1 and PAS3. Data are average values of three determinations.

Enzyme	Substrate	K_m (µM)	V_{max} (nkat/mg)	K_{cat} (min^{-1})	K_{cat}/K_m ($S^{-1}M^{-1}$)
PAS1	Cinnamoyl-CoA	7.15	9.35	2.94	6853
	p-Coumaroyl-CoA	9.21	6.26	1.99	3601
	Malonyl-CoA	28.56			
PAS3	Cinnamoyl-CoA	6.03	7.89	2.51	6937
	p-Coumaroyl-CoA	7.78	5.16	1.64	3513
	Malonyl-CoA	20.16			

The kinetic analysis revealed that both PAS1 and PAS3 accepted *p*-coumaroyl-CoA with less efficiency than cinnamoyl-CoA. The K_m values of PAS1 and PAS3 for *p*-coumaroyl-CoA were about 1.3-fold higher and thus the affinity lower. Similarly, the V_{max} and K_{cat} values for *p*-coumaroyl-CoA were lower by approx. one third. As a consequence, the K_{cat}/K_m values for *p*-coumaroyl-CoA were only about 50% of those for cinnamoyl-CoA. PAS3 exhibited higher affinity for the substrates than PAS1 but had lower K_{cat} values, resulting in similar K_{cat}/K_m values for the substrates.

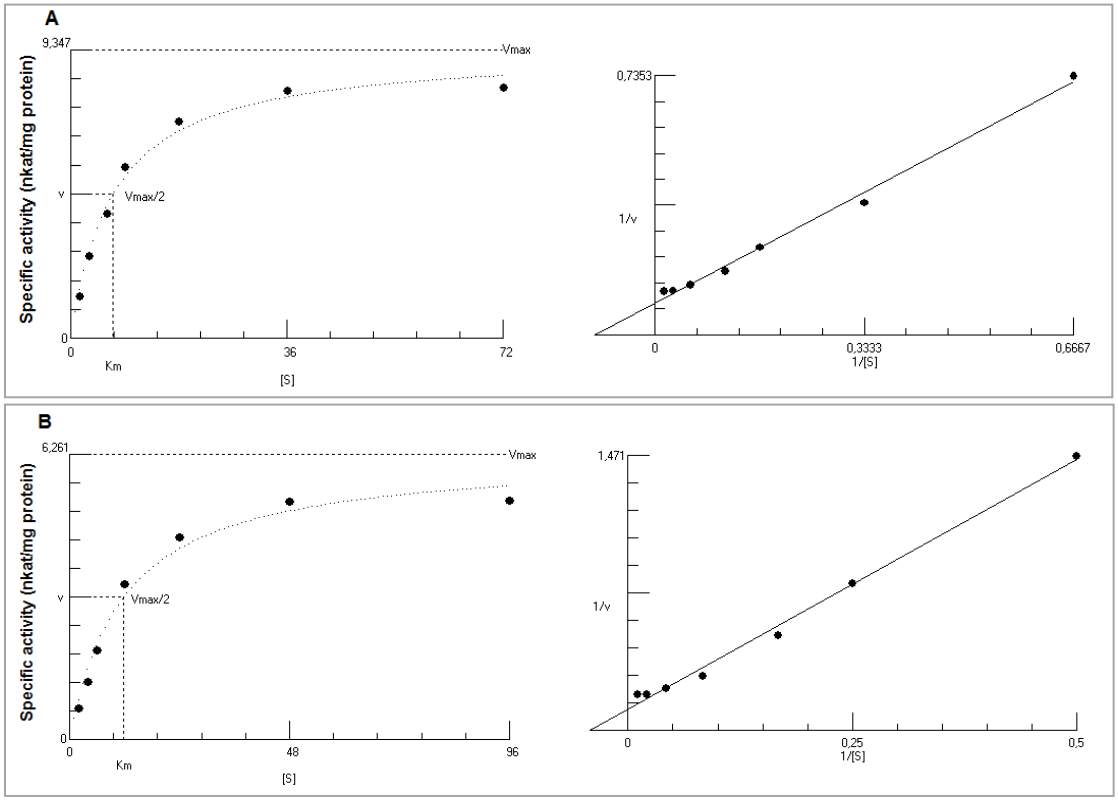


Fig. 49. Dependence of PAS1 activity on the substrate concentration and determination of the K_m value *via* Lineweaver-Burk plot. A. Cinnamoyl-CoA, B. *p*-Coumaroyl-CoA.

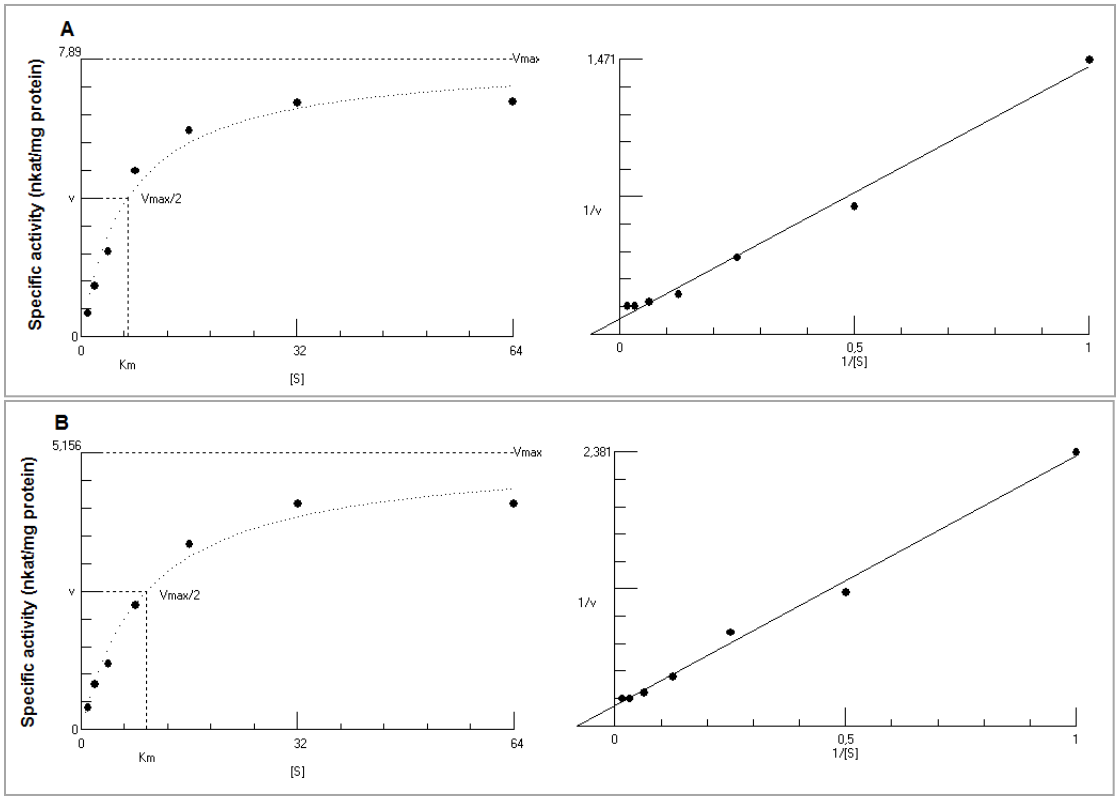


Fig. 50. Dependence of PAS3 activity on the substrate concentration and determination of the K_m value *via* Lineweaver-Burk plot. A. Cinnamoyl-CoA, B. *p*-Coumaroyl-CoA.

4.5.2 Biochemical characterization of C.CHSs

4.5.2.1 Determination of pH optima

The activity of the C.CHSs was determined between pH 3.0 and 10.0 in the standard assay. The dependence of enzyme activities on pH is shown in Fig. 51. Although the pH optima were similar, there were distinct properties when compared to each other. The maximum catalytic activity and stability of all C.CHSs were at pH 7.0-7.5 except for C.CHS3. This enzyme was more active under alkaline conditions and the optimum pH was at 8.0. The data shown are mean values of three experiments. Consequently, all subsequent tests were performed at pH 7.0-8.0 in potassium phosphate buffer. At this optimum pH, another series of incubations were performed at different temperatures to determine the optimum temperature

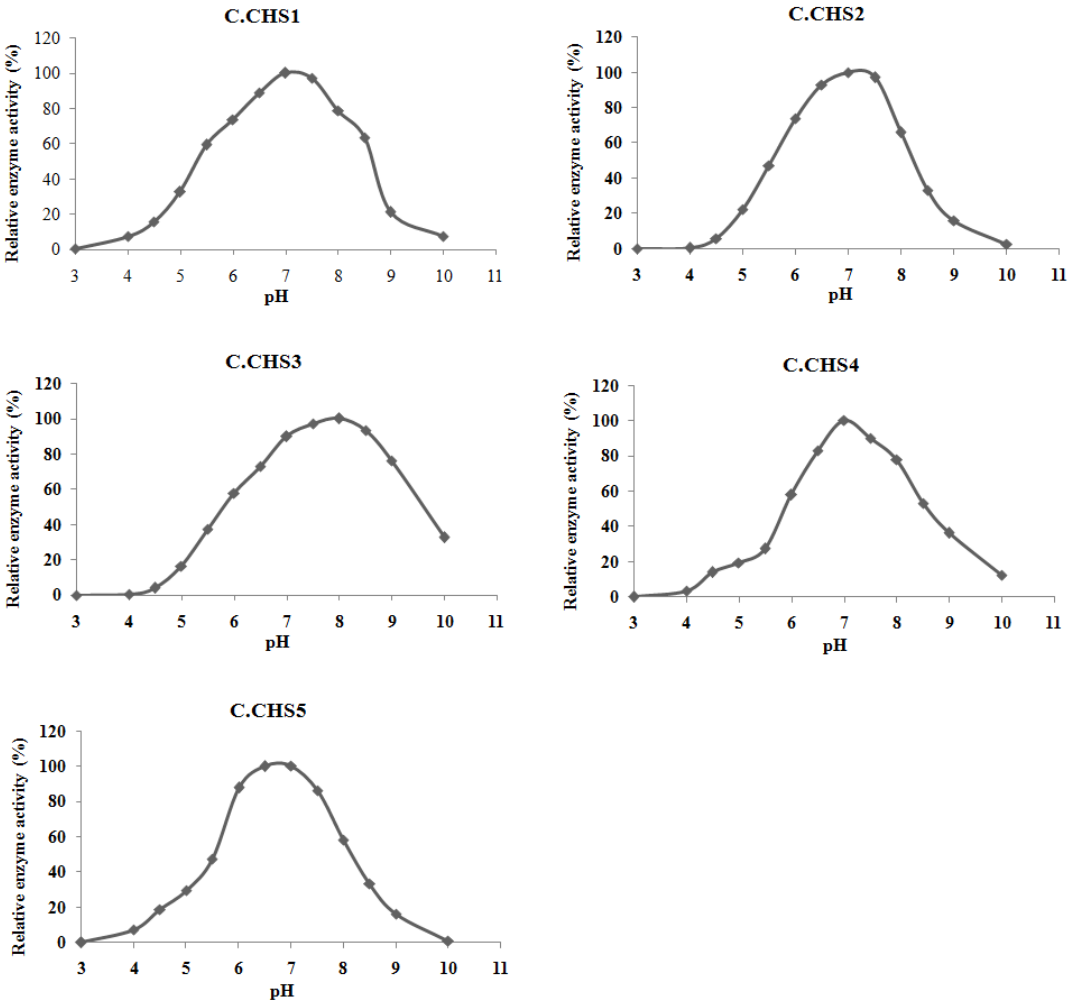


Fig. 51. Effect of pH on the activities of the C.CHSs

4.5.2.2 Temperature optima

Standard assays were performed at 5°C intervals from 20 to 55°C and average values of three experiments were calculated. Each enzyme has a temperature range in which a maximal rate of reaction is achieved and which is known as the temperature optimum of the enzyme. The temperature optima of all C.CHSs were at 35°C, their activities being high between 30 and 40°C. At incubation temperatures above 40°C, the proteins began to denature, lost their functions and became inactive to operate properly (Fig. 52).

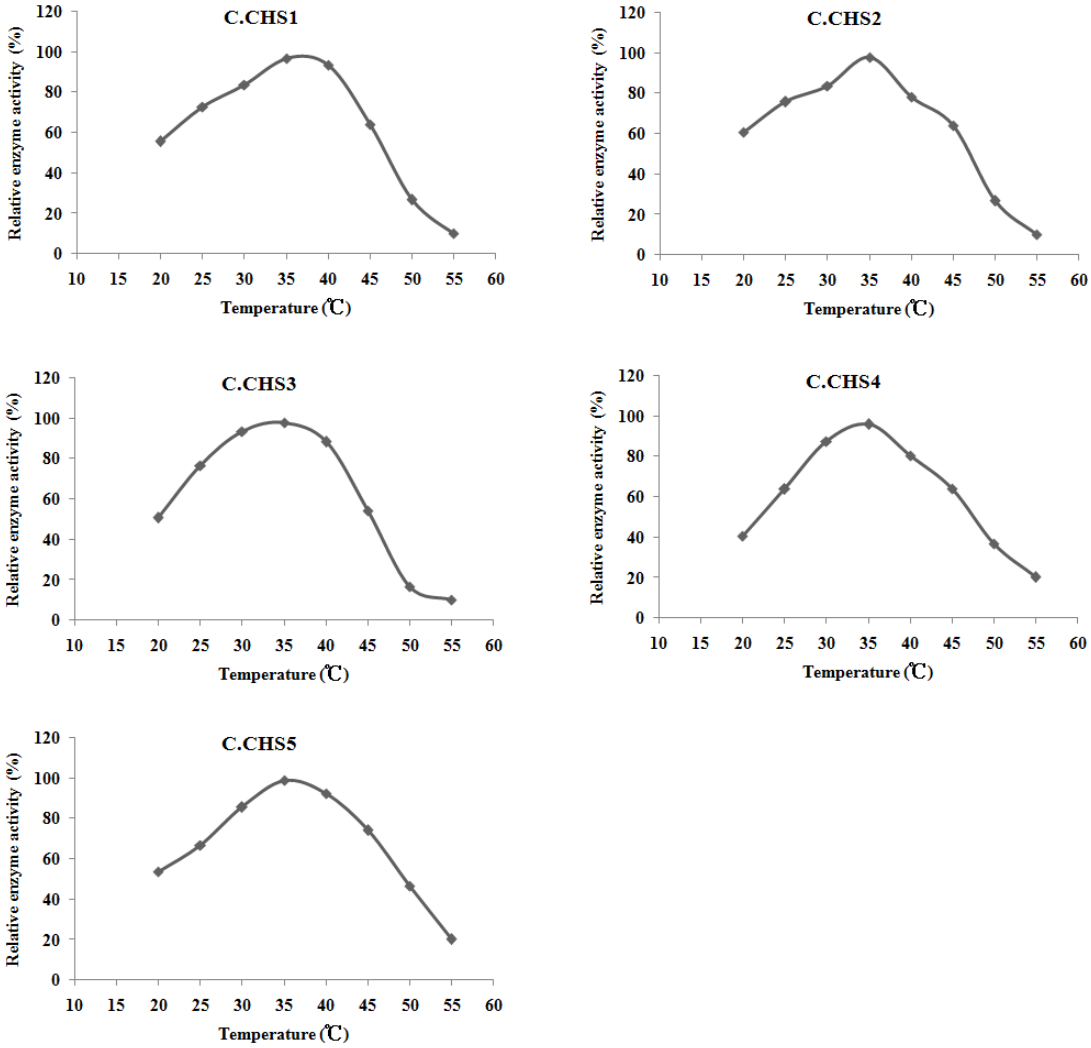


Fig. 52. Effect of temperature on the activities of the C.CHSs .

4.5.2.3 Linearity with protein amount and incubation time

The amount of enzymatically formed pinocembrin was determined as a function of the protein amount in the standard assay.The accumulation of the product pinocembrin was linear up to 15 min of incubation time and up to 8 µg of both C.CHS1 and C.CHS2 protein per assay (Fig. 53, 54). For C.CHS3, product fromation was linear with a protein concentration up to 10 µg and up to 30 min of incubation time in the standard assay. In case of C.CHS4, the product accumulated linearly up to 20 min of incubation time and up to 12 µg of protein per assay. Finally, for C.CHS5 accumulation of product was linear up to 20 min of incubation time and up to 6 µg of protein per assay.

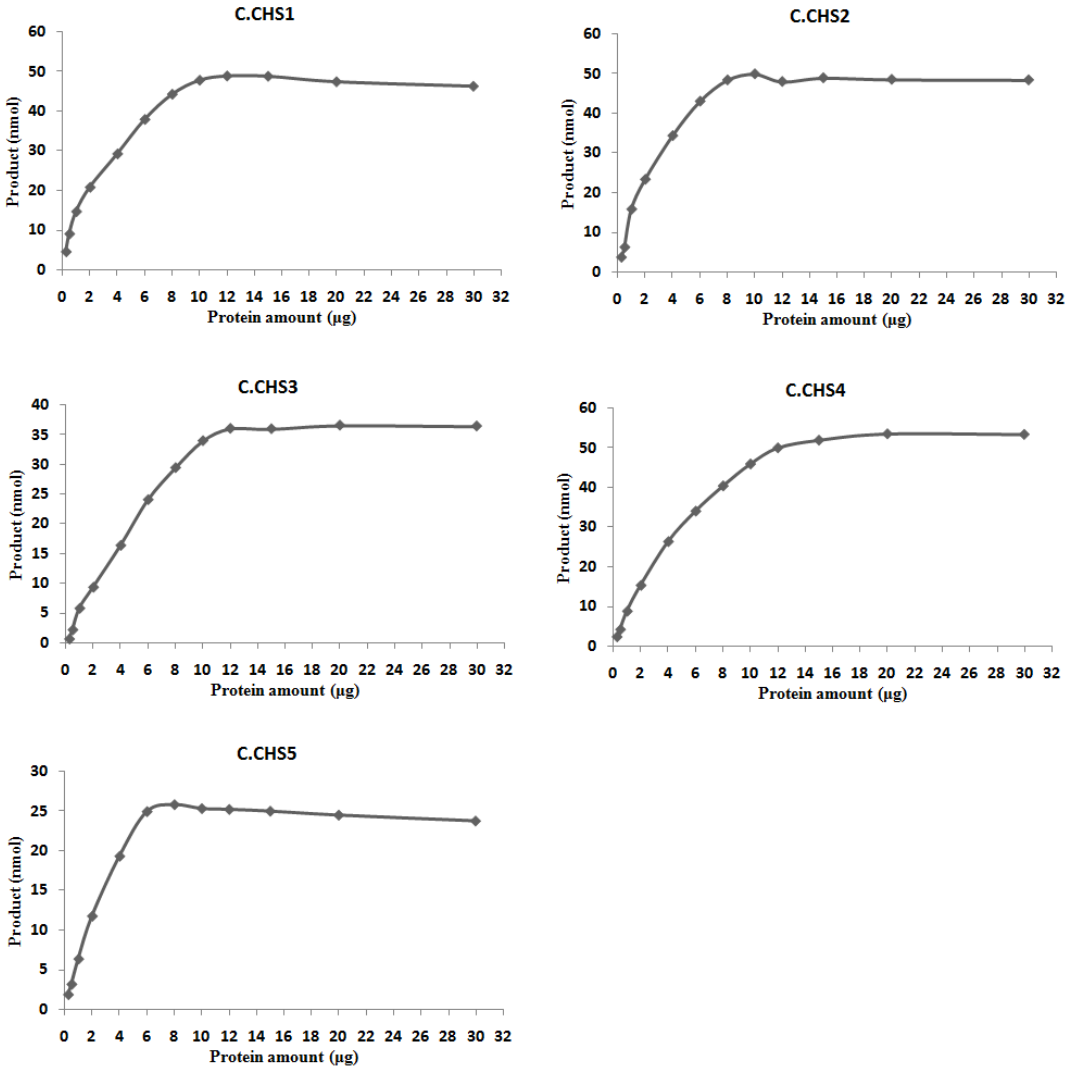


Fig. 53. Effect of protein amount on the activities of the C.CHSs.

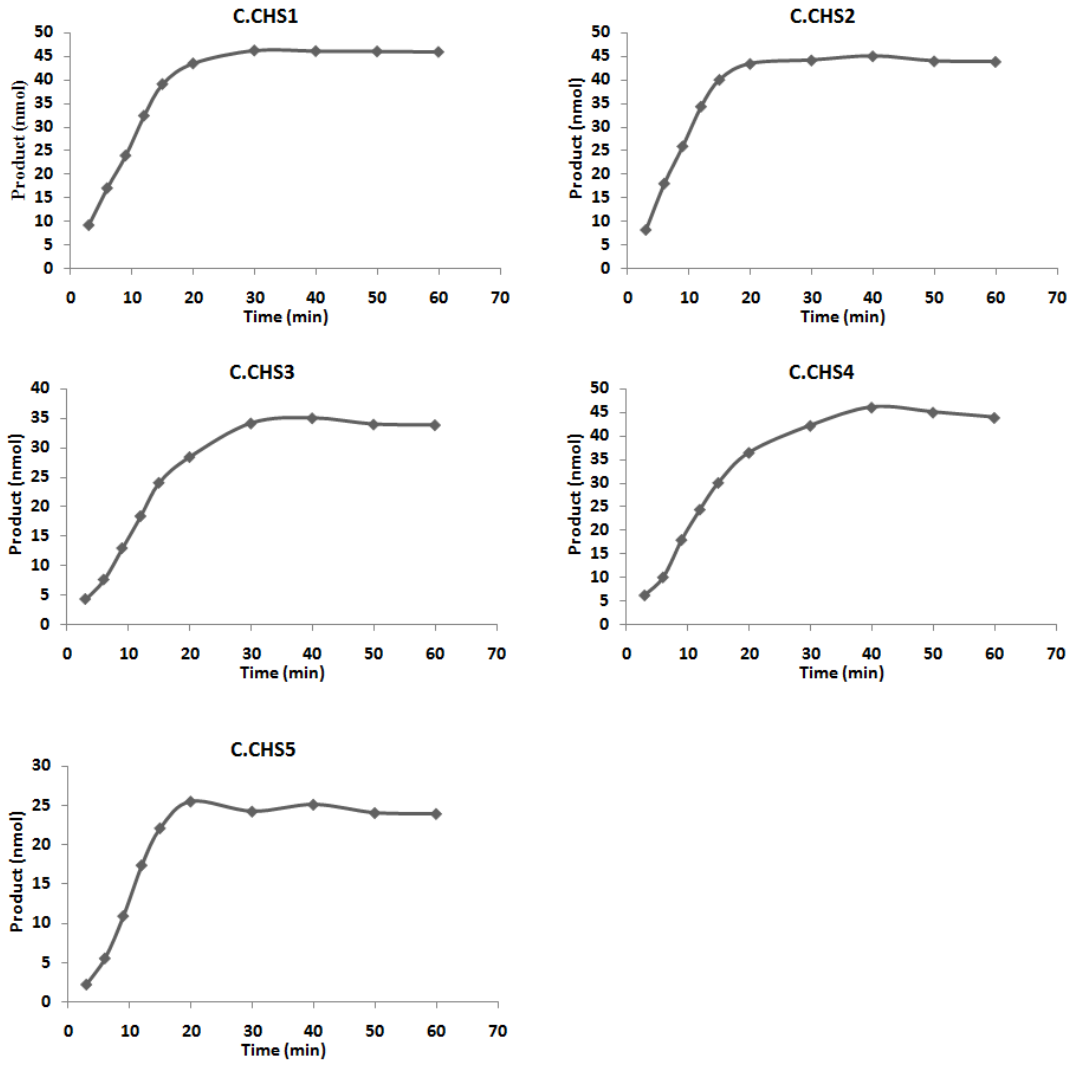


Fig. 54. Effect of incubation time on the activities of the C.CHSs.

4.5.2.4 Determination of substrate specificities

All C.CHSs were tested under the respective optimum assay conditions with a series of potential substrates. The C.CHSs were incubated in the presence of malonyl-CoA as extender with cinnamoyl-CoA, *p*-coumaroyl-CoA, dihydro-*p*-coumaroyl-CoA, dihydro-cinnamoyl-CoA, acetyl-CoA, benzoyl-CoA, *m*-coumaroyl-CoA, *o*-coumaroyl-CoA and salicoyl-CoA. All C.CHSs had distinct properties when incubated with the different substrates. Different from most published CHSs, the recombinant enzymes encoded by C.CHS1, C.CHS2 and C.CHS4 preferred cinnamoyl-CoA as starter substrate instead of *p*-coumaroyl-CoA (Table 13). Experiments with these substrates showed that the product amounts were about 30% higher than those with *p*-coumaroyl-CoA. These CHSs were likely to be responsible for the biosynthesis of related flavonoids in *C. cajan*, which possess an unsubstituted B ring with cinnamoyl-CoA as precursor. Conversely, C.CHS3 and C.CHS5 exhibited highest activity with *p*-coumaroyl-CoA and cinnamoyl-CoA was only the second best substrate

(70% relative activity). They should be responsible for the biosynthesis of other kinds of flavonoids, which possess a substituted B ring in this plant. Additionally, all C.CHSs also accepted benzoyl-CoA but produced exclusively benzoyldiacetic acid lactone. None of the enzymes accepeted acetyl-CoA, *m*-coumaroyl-CoA, *o*-coumaroyl-CoA and salicoyl-CoA. The identities of all the reaction products were verified by HPLC analysis in comparison with authentic reference compounds and by LC-MS analysis. Product amounts were quantified by radio activity detector-coupled HPLC analysis.

Table 13. Substrate specificities of C.CHSs

Starter substrate	**Product (% of maximum each)**				
	C.CHS1	**C.CHS2**	**C.CHS3**	**C.CHS4**	**C.CHS5**
cinnamoyl-CoA	100	100	73	100	67.5
p-coumaroyl-CoA	71.5	73.5	100	70.8	100
dihydro-cinnamoyl-CoA	52.3	50.6	31.5	47.2	33.9
dihydro-*p*-coumaroyl-CoA	37.5	43.6	56.8	32.5	55.2
m-coumaroyl-CoA	0	0	0	0	0
o-coumaroyl-CoA	0	0	0	0	0
benzoyl-CoA	35.3	40.7	48.5	36.2	47.7
acetyl-CoA	0	0	0	0	0
salicoyl-CoA	0	0	0	0	0

4.5.2.5 Determination of kinetic parameters

The kinetic properties of the C.CHSs were determined under the optimum conditions. The summaries of the kinetic data for all the enzymes are presented in Table 14. Each assay contained 2 µg protein and the incubation period was limited to 10 min. Besides the K_m values, the K_{cat} and K_{cat}/K_m values were determined on the basis of the Lineweaver-Burk plots at saturating concentrations of both substrates.

Kinetic analysis revealed that C.CHS1, C.CHS2 and C.CHS4 exhibited higher turnover rates (V_{max} and K_{cat}) for cinnamoyl-CoA than for *p*-coumaroyl-CoA and in addition their affinities (K_m) for cinnamoyl-CoA were somewhat higher than for *p*-coumaroyl-CoA. The K_{cat}/K_m values demonstrated that these three enzymes accepted *p*-coumaroyl-CoA with only 65-70% efficiency when compared with cinnamoyl-CoA. C.CHS2 was the enzyme that was catalytically most efficient with cinnamoyl-CoA. Incontrast, it was remarkable that C.CHS3 and C.CHS5 preferred *p*-coumaroyl-CoA as a starter molecule, as reflected by the K_{cat}/K_m values. C.CHS5 has the highest affinity and catalytic efficiency towards *p*-coumaroyl-CoA.

Table 14. Kinetic parameters of C.CHSs. Data are average values of three determinations.

Enzyme	Substrate	K_m (µM)	V_{max} (nkat/mg)	K_{cat} (min^{-1})	K_{cat}/K_m ($S^{-1}M^{-1}$)
C.CHS1	Cinnamoyl-CoA	6.27	17.47	5.58	14832
	p-Coumaroyl-CoA	7.28	12.33	3.94	9020
	Malonyl-CoA	32.13			
C.CHS2	Cinnamoyl-CoA	5.15	19.56	6.24	20194
	p-Coumaroyl-CoA	6.34	15.78	5.03	13240
	Malonyl-CoA	25.26			
C.CHS3	Cinnamoyl-CoA	10.75	9.47	3.02	4682
	p-Coumaroyl-CoA	9.37	16.33	5.21	9267
	Malonyl-CoA	26.88			
C.CHS4	Cinnamoyl-CoA	8.53	13.87	4.42	8636
	p-Coumaroyl-CoA	9.26	9.65	3.08	5543
	Malonyl-CoA	35.59			
C.CHS5	Cinnamoyl-CoA	5.08	10.16	3.24	10630
	p-Coumaroyl-CoA	4.78	13.85	4.42	15411
	Malonyl-CoA	17.23			

5. Molecular cloning of cDNAs encoding potential polyketide cyclases involved in stilbenecarboxylate biosynthesis in *C. cajan* leaves

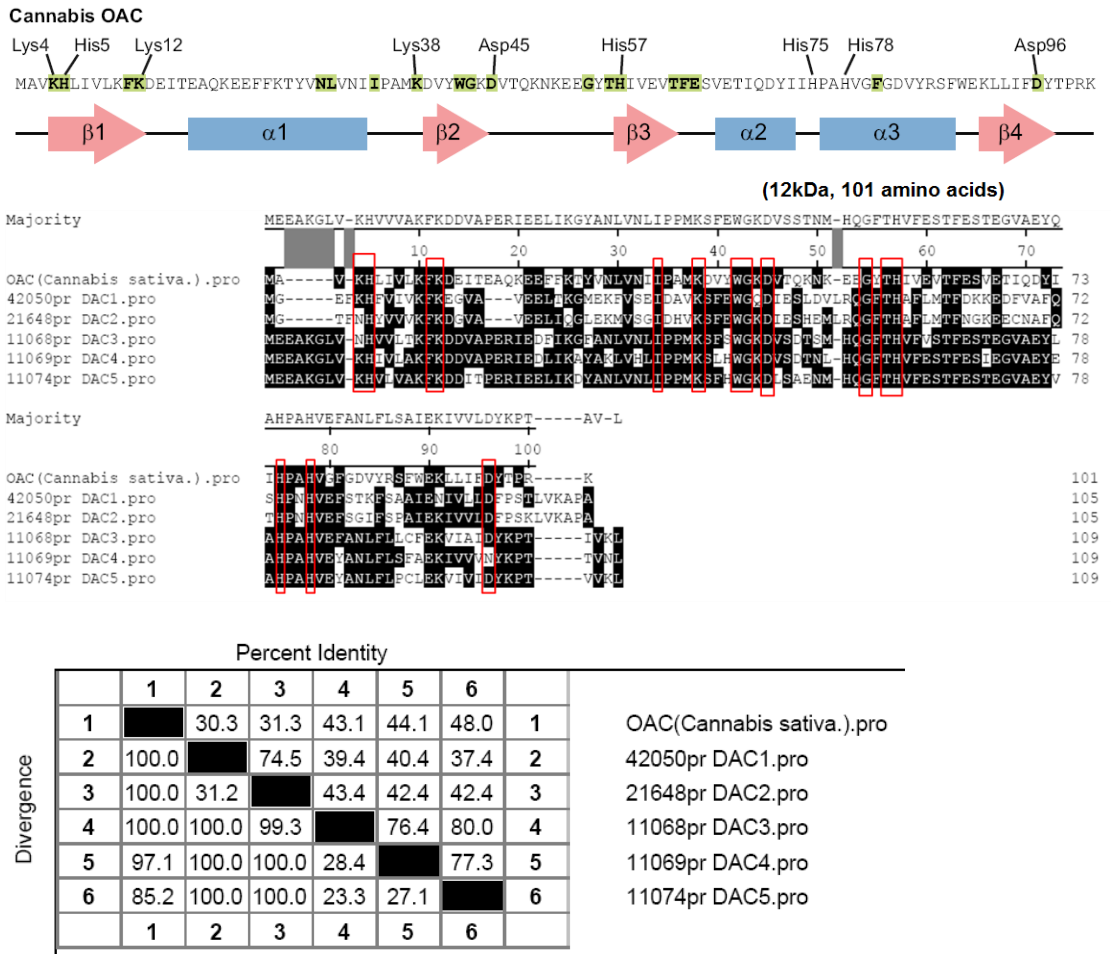
As mentioned above, the independent findings from different approaches suggested that there are two types of polyketide synthase within the CHS-type protein family in *C. cajan*: C.CHSs participate in flavonoid biosynthesis, and PASs performed two or three condensations with malonyl-CoA and released the corresponding lactones but neither chalcone or stilbene nor stilbenecarboxylate. This means that the PKS which is responsible for stilbenecarboxylate biosynthesis in this plant has not yet been identified at the enzyme or the gene level. On the other hand, this plant contains a large amount of stilbenecarboxylates and their derivatives (e.g. cajaninstilbene acid), which are most likely derived from cinnamoyl-CoA according to their structure characteristics (Fig. 10). Meanwhile, experiments of substrate specificities had shown that the recombinant PAS1 and PAS3 also preferred cinnamoyl-CoA as starter substrate. Therefore, it is reasonable to assume that the PASs probably provide only such a precursor and that CITAL was probably a derailment product of the PASs due to the absence of the subsequent reactions. The correct cyclization of the linear polyketide intermediate might need the help of a specific cyclase, which coacts with PAS and finally accounts for the biosynthesis of stilbenecarboxylate. This assumption has been proved to be true by the recently published work on olivetolic acid biosynthesis in *Cannabis sativa* (Gagne et al. 2012). These authors showed that a type III PKS from cannabis trichomes requires the presence of a polyketide cyclase enzyme, olivetolic acid cyclase (OAC), which catalyzes a C2–C7 intramolecular aldol condensation with carboxylate retention to form olivetolic acid. OAC is a dimeric α+β barrel (DABB) protein that is structurally similar to polyketide

cyclases from *Streptomyces* species. According to these results, we also attempted to find a stilbenecarboxylate cyclase from *C. cajan*.

A total of 11 candidate genes were detected in the *C. cajan* genome using the polyketide cyclase sequences that have been mentioned in the studies of the Page group as blasting probes. The 11 candidates can be subdivided in three kinds of polyketide cyclases based on their predicted function, which are called CHI-like, DABB, and Betv1-like protein (Table 15). A chalcone isomerase (CHI)-like protein was selected based on the catalytic relationship of CHS with CHI (Jez et al., 2000, Burbulis et al., 1999). The second candidate was a member of the dimeric α+β barrel (DABB) protein superfamily with similarity to stress-responsive proteins in plants (Wang et al. 2002; Park et al. 2007; Lee et al., 2008). The presence of this protein was intriguing because DABB proteins act as polyketide cyclases (e.g., TcmI cyclase) in *Streptomyces* species (Thompson et al., 2004), although the bacterial cyclases show low sequence similarity to plant DABB proteins. The third candidate was a Betv1-like protein in the same protein family as the *Streptomyces* TcmN ARO/CYC polyketide cyclase (Ames et al. 2008). Several Betv1-protein family members function as enzymes in plant natural product biosynthesis (Radauer et al., 2008). The published olivetolic acid (OAC) cyclase is just a dimeric α+β barrel (DABB) protein from *Cannabis sativa.* So, alignment of DABB protein sequences from *C. cajan* with OAC (*Cannabis sativa.*) was carried out and is presented in Fig. 55. This alignment shows that five DABB candidates share 30-48% identity with OAC and all probable active sites of OAC are present in the sequences of these proteins. Consequently, these selected proteins were heterologously expressed and their function was studied in combination with the PASs of *C. cajan* in the biosynthesis of stilbenecarboxylate. The corresponding primer sets are mentioned under II.1.5.2.

Table 15. Polyketide cyclase candidates identified in the *C. cajan* genome.

DABB protein (DAC) Stress responsive dimeric α+β barrel (DABB) domain family	*C. cajan*_42050 (**DAC1**)
	*C. cajan*_21648 (**DAC2**)
	*C. cajan*_11068 (**DAC3**)
	*C. cajan*_11069 (**DAC4**)
	*C. cajan*_11074 (**DAC5**)
Betv1-like protein (BEC) Pathogenesis-related protein Betv1 family	*C. cajan*_31770 **(BEC1)**
	*C. cajan*_47832 **(BEC2)**
	*C. cajan*_42730 **(BEC3)**
	*C. cajan*_04110 **(BEC4)**
	*C. cajan*_40987 **(BEC5)**
CHI-like protein (CHC) Chalcone-flavanone isomerase family	*C. cajan*09652 **(CHC1)**



	1	2	3	4	5	6	
1		30.3	31.3	43.1	44.1	48.0	1
2	100.0		74.5	39.4	40.4	37.4	2
3	100.0	31.2		43.4	42.4	42.4	3
4	100.0	100.0	99.3		76.4	80.0	4
5	97.1	100.0	100.0	28.4		77.3	5
6	85.2	100.0	100.0	23.3	27.1		6
	1	2	3	4	5	6	

Fig. 54. Alignment of DABB protein sequences from *C. cajan* with OAC (*Cannabis sativa*). Red boxes show probable active site residues of OAC.

5.1 Amplification of cDNAs encoding potential polyketide cyclases

A cDNA pool was prepared by reverse transcription of RNA isolated from *C. cajan* leaves and used as a template for PCR reaction (II.2.3.3). The obtained cDNA was examined for its quality by checking the successful expression of the 18s rRNA. Full length genes were amplified using gene-specific primers basedon the obtained open-reading frame (ORF) sequences. The primer sets used for amplification of the candidates are mentioned under II.1.5.2. The cDNAs were amplified using proof-reading Phusion DNA polymerase Hot start II and the PCR program under II.2.3.5. In all of candidate sequences, the restriction site *Nhe* I was integrated in the forward primers, while the restriction site *Kpn* I was integrated in the reverse primers. With standard PCR programme and an annealing temperature of 63°C, the primerswere designed to yield five products for DAC1-5, five products for BEC1-5 and one product for CHC1 (Fig. 56). All of them were used in the following work. Their nucleotide sequences are included in the appendix.

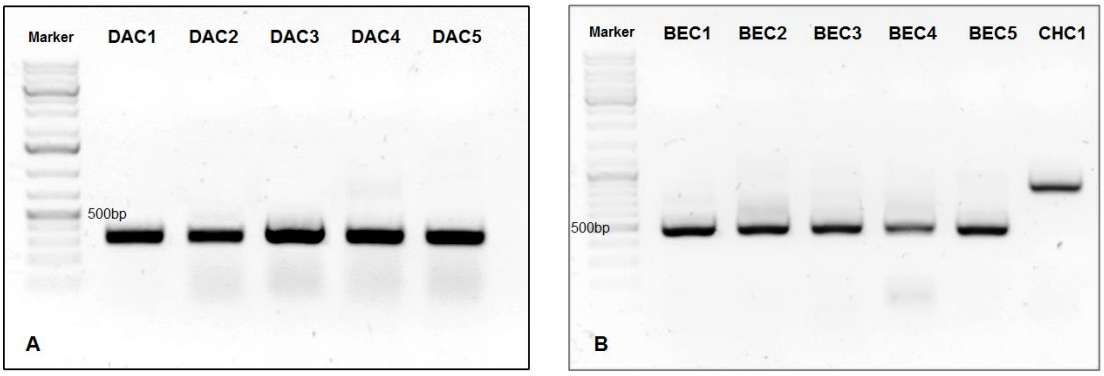


Fig. 56. Agarose gel electrophoresis of PCR products for polyketide cyclases from *C. cajan.* A. PCR products for DACs, B. PCR products for BECs and CHC.

5.2 Heterologous expression in *E. coli* and enzyme purification

The PCR products from the last step were purified, digested and ligated to the pRSET B expression vector between the corresponding restriction sites (II.2.3.8, II.2.3.9). The ligation products were introduced in *E. coli* DH5α (II.2.3.10) to obtain a large amount of the plasmids, followed by plasmid isolation (II.2.3.11). Confirmation of successful ligation was carried out by restriction digestion analysis. The constructed plasmids were sequenced to ensure the presence of the right insert, and no frame shifts were detected. Plasmids were transferred in *E. coli* strain BL21 (DE3) pLysS for carrying out heteologuos expression (II.2.3.12). They were functionally expressed as 6xHis-tagged proteins. Following protein extraction and purification, the efficiency of overexpression and confirmation of their production as soluble proteins were examined by SDS-PAGE (Fig. 57). The Ni-NTA system was used for the purification of the 6xHis-tagged cyclases. The proteins induced upon addition of IPTG constituted major bands in the crude soluble protein. Finally, they were eluted from the Ni-NTA matrix. The yield of pure protein was determined by the Bradford method and a certain amount of purified protein was used in the following tests for enzymatic activity.

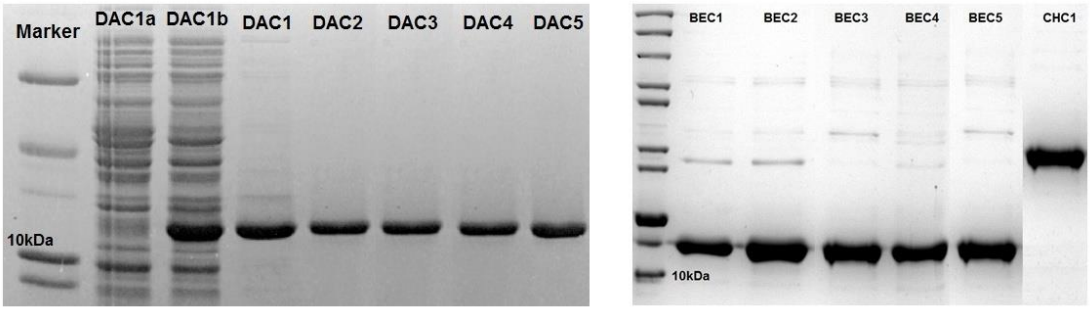


Fig. 57: SDS-PAGE of over-expressed putative polyketide cyclase proteins
DAC1a: Crude protein DAC1 before induction with IPTG (pellet)
DAC1b: Crude protein DAC1 after induction with IPTG (pellet)
DAC1-5, BEC1-5 and CHC1: Purified protein (≈ 5µg each)

5.3 Test for enzymatic activity

The functions of the purified fusion proteins were investigated using an in vitro enzyme reaction. The procedure of enzyme assay is described in II.2.2.4. PAS (10 μg), DAC (7 μg), BEC (7 μg), and CHC (7 μg) were added individually or in combination. Reaction mixtures were incubated at 30°C for 60 min. The assays were carried out with cinnamoyl-CoA and malonyl-CoA as the substrates. Confirmation of the enzyme activity was carried out doing parallel incubations containing only PAS or cyclase. After termination by acidification, the reaction products were extracted by ethyl acetate and the extracts then analyzed by HPLC.
All the enzyme assays failed to demonstrate activity for stilbenecarboxylate biosynthesis. No additional peaks or increases in the original peaks were detected in the co-incubations. All the mixed enzyme assays produced only the corresponding lactones, which are also observed in the assays containing PAS only. Different protein concentrations, different reaction times and different temperatures of incubation were also tried inthe assays, but no additional new product was detected (Fig. 58).
Despite our considerable efforts, we failed to detect the STCS activity of PAS in the presence of a putative polyketide cyclase. There are several possible explanations for this negative result. One is the presence of other yet unidentified polyketide cyclases specific for stilbenecarboxylate biosynthesis, which were not be obtained by PCRs under the conditions employed in this study. The second is that the identitied PASs in our study are not the related PKSs responsible for stilbenecarboxylate biosynthesis in this plant. The polyketide cyclases generally had no intrinsic PKS activity and only function in combination with PKS to help polyketide assembly. The third is that the presence of a specific cyclase, which coacts with the required PKS to produce stilbenecarboxylate, must be carried out under distinct yet unknown incubation conditions. The close association with the cyclase protein and the correct cyclization of the linear polyketide intermediate with carboxylate retention might need the help of yet unknown factors. The *in vitro* incubation may lack some conditions that are necessary for the correct function of the protein. Further investigations are required to uncover the actual mechanism of stilbenecarboxylate biosynthesis in this plant.

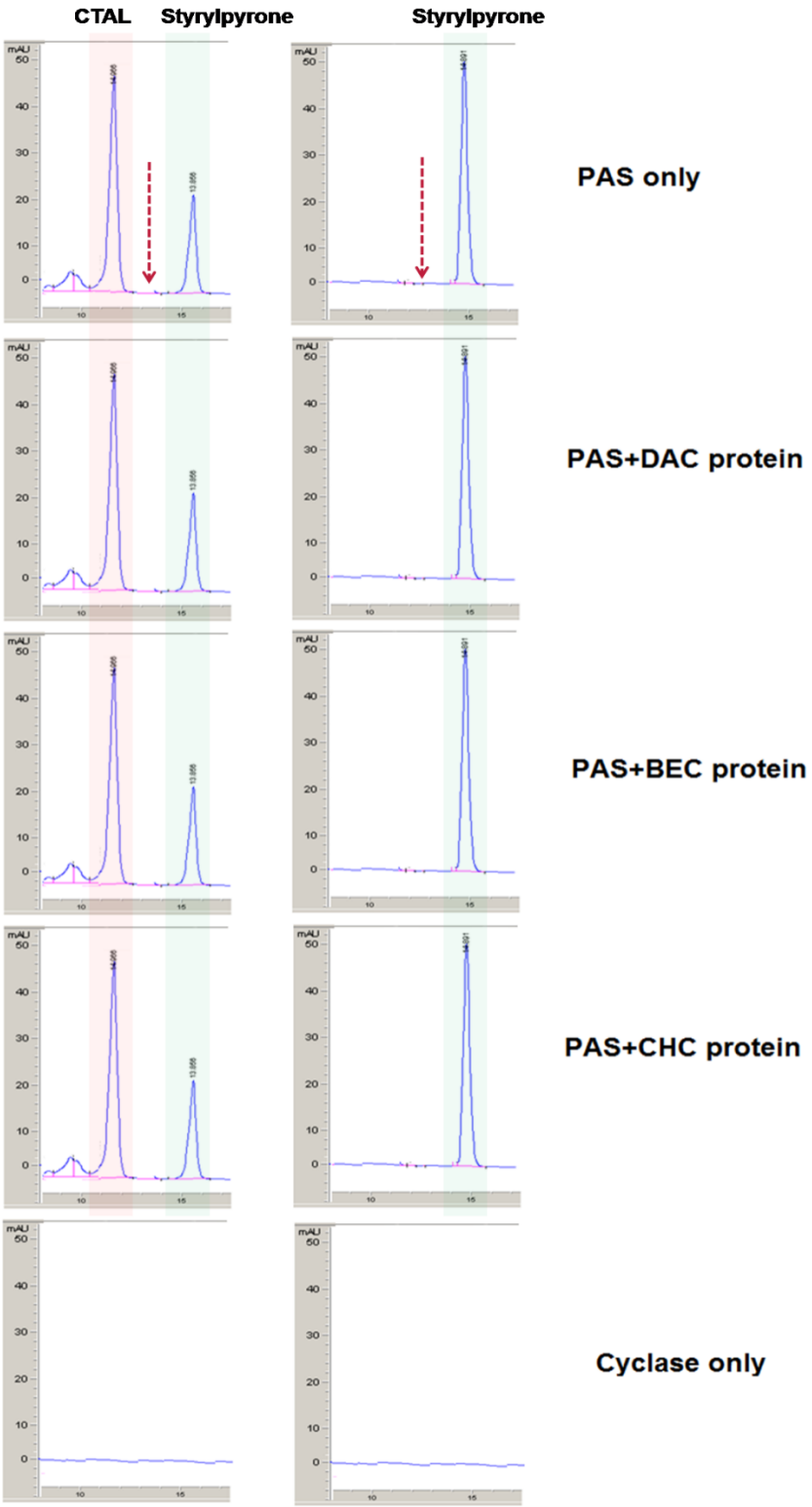


Fig. 58. HPLC analysis of polyketide products in PAS assays containing in addition different cyclases. Incubation with only PAS or cyclase served as controls. Arrow: expected peak for stilbenecarboxylate

6. Organ-specific expression of type III polyketide synthase genes in *C. cajan*

As described above, stilbenecarboxylate formation was not detected in PKS assays. However, it was confirmed before that the plant *C. cajan*, from which these PKS cDNAs were cloned, contains STCS derivatives, such as cajaninstilbene acid. The distribution of the metabolic products thus should be correlated with the expression patterns of the PKS genes, as determined by transcript analysis in total RNA isolated from different plant organs at various growth periods. Expression of the PAS and C.CHS genes was analyzed by RT-PCR using gene-specific primers after optimization of the cycle number. Actin served as a control to ensure equality of RNA levels used (Fig. 59).

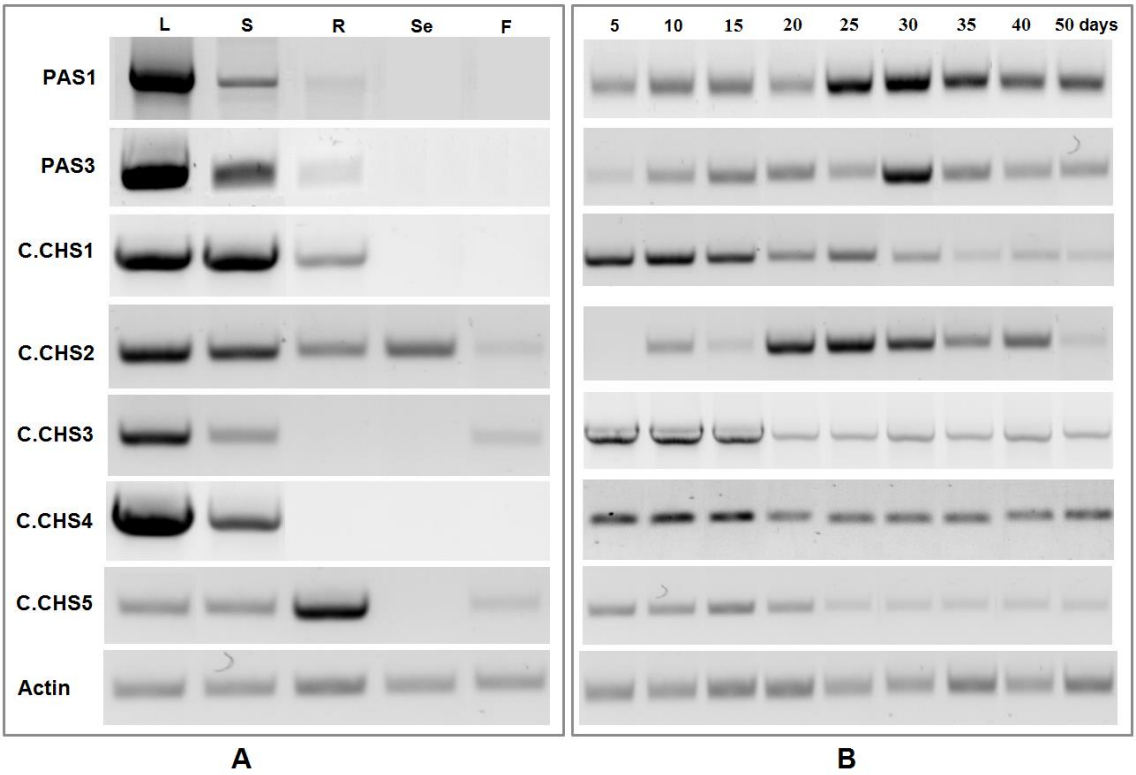


Fig. 59. Semiquantitative RT-PCR analysis of PAS and C.CHS gene expression.
A: For organ-specific expression, total RNA was prepared from L: leaves, S: stems, R: roots, Se: seeds and F: flowers.
B: For expression in leaves at various growth periods, leaves were collected from the seedlings until day 50 after germination and used for extraction of total RNA

For organ-specific expression of PKSs, semiquantitative RT-PCR analysis revealed that the accumulation of all PKSs transcripts was more abundant in leaves and stems than in roots, seeds and flowers, except for C.CHS5. The maximum C.CHS5 transcript level was observed in roots and the transcript was only present at a low level in leaves. High expression levels of PAS1 and PAS3 were detected in leaves, where the STCS derivative levels were also the highest. In seeds and flowers, these PKS transcripts were present at

low levels.

The expression profiles of the investigated genes in leaves at various growth periods were interesting. PAS1 expression paralleled that of PAS3. Their transcripts started to accumulate at the onset of seedling germination, reached a maximum after 25 to 30 days, whereafter the transcript levels gradually decreased. In contrast to the PASs, the expression patterns of the C.CHS were distinct. All C.CHS transcripts were detectable during the entire growth period of the leaves. Most of them reached a maximum at the beginning of seedling growth, which suggested that the accumulation of these transcripts was more abundant in young leaves. However, C.CHS2 performed in a different way because a high expression level was detected only after 20 days. Thereafter, the transcript level started to decline and stabilized at a relatively low level over the test days.

IV. Discussion

1. Type III polyketide synthases in *Cajanus cajan*

C. cajan not only contains stilbenecarboxylate, but also other kinds of secondary metabolites, which exhibit excellent biological activities such as flavonoids, stilbenes and isocoumarins. During the last few decades much attention has been given to the isolation and identification of these compounds. Extensive studies have been carried out regarding the chemistry of *C. cajan* and considerable progress has been achieved regarding its biological activities and medicinal applications. However, there are not yet any data on the biosynthesis of secondary metabolites in *C. cajan*. In the four above-mentioned secondary metabolite groups (flavonoids, stilbenoids, isocoumarins and stilbenecarboxylates), enzymes of the type III polyketide synthase group are involved in the biosynthesis of their initial skeletons. Chalcone synthase (CHS), stilbene synthase (STS) and related proteins are key enzymes in the biosynthesis of flavonoids and stilbenes, respectively, and they have been extensively studied in several plants (Austin and Noel, 2003; Kuhn et al., 2011; Ono et al., 2006; Saslowsky et al., 2005). The detection of stilbenecarboxylate synthase (STCS) activity and its molecular characterization (Austin et al., 2004) has not ended the debate about the origin of the terminal carboxyl group of stilbenecarboxylate. Precursor feeding studies and mechanistic rationalization suggest that stilbenecarboxylates are synthesized by a plant type III polyketide synthase; however, the enzyme activity leading to retention of the carboxyl moiety in a stilbene backbone has not yet been completely elucidated. *C. Cajan* contains a good amount of stilbenecarboxylate such as cajaninstilbene acid, isocajaninstilbene acid and 4-*O*-methylpinosylvic acid. Even more interesting is that a series of products which belong to stilbenes and isocoumarins may associate with the formation of stilbenecarboxylate (Fig. 60). They are likely to have derived skeletons, which have similar substitution patterns. Their simultaneous production in the plant was detected in extensive studies and they might be involved in the postulated biosynthetic pathway (Fig. 61).

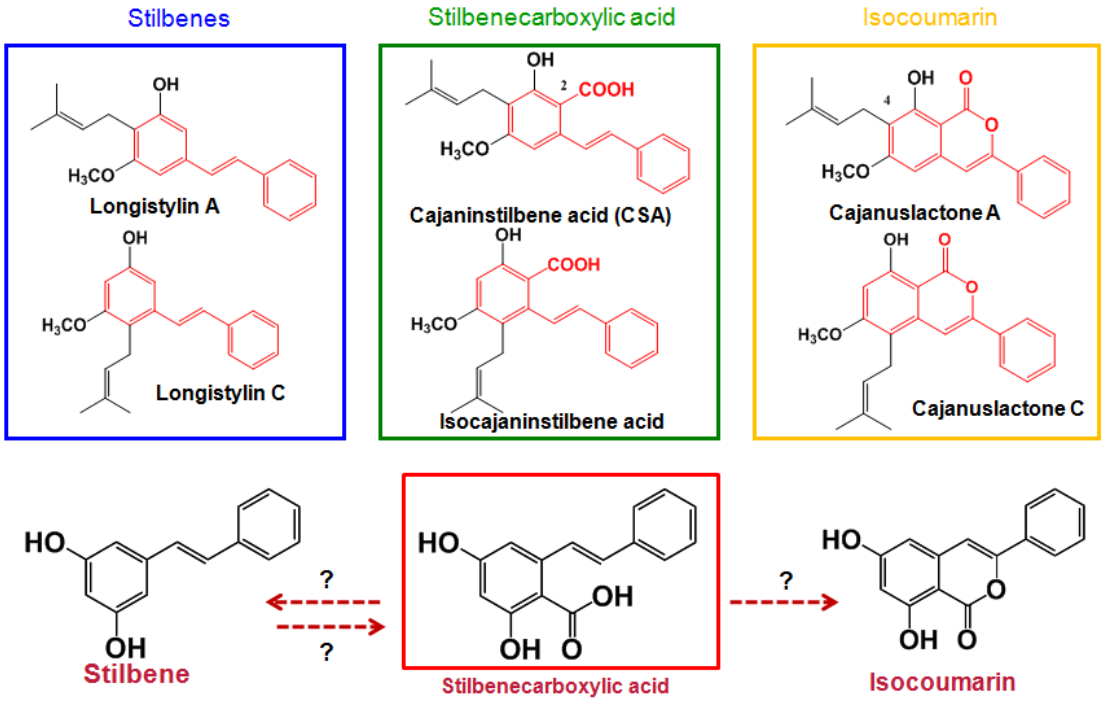


Fig. 60. Stilbenecarboxylate and derivatives in *C. cajan* and the proposed biosynthetic pathway. The blue, green and yellow boxes show the classes of natural products identified in the plant.

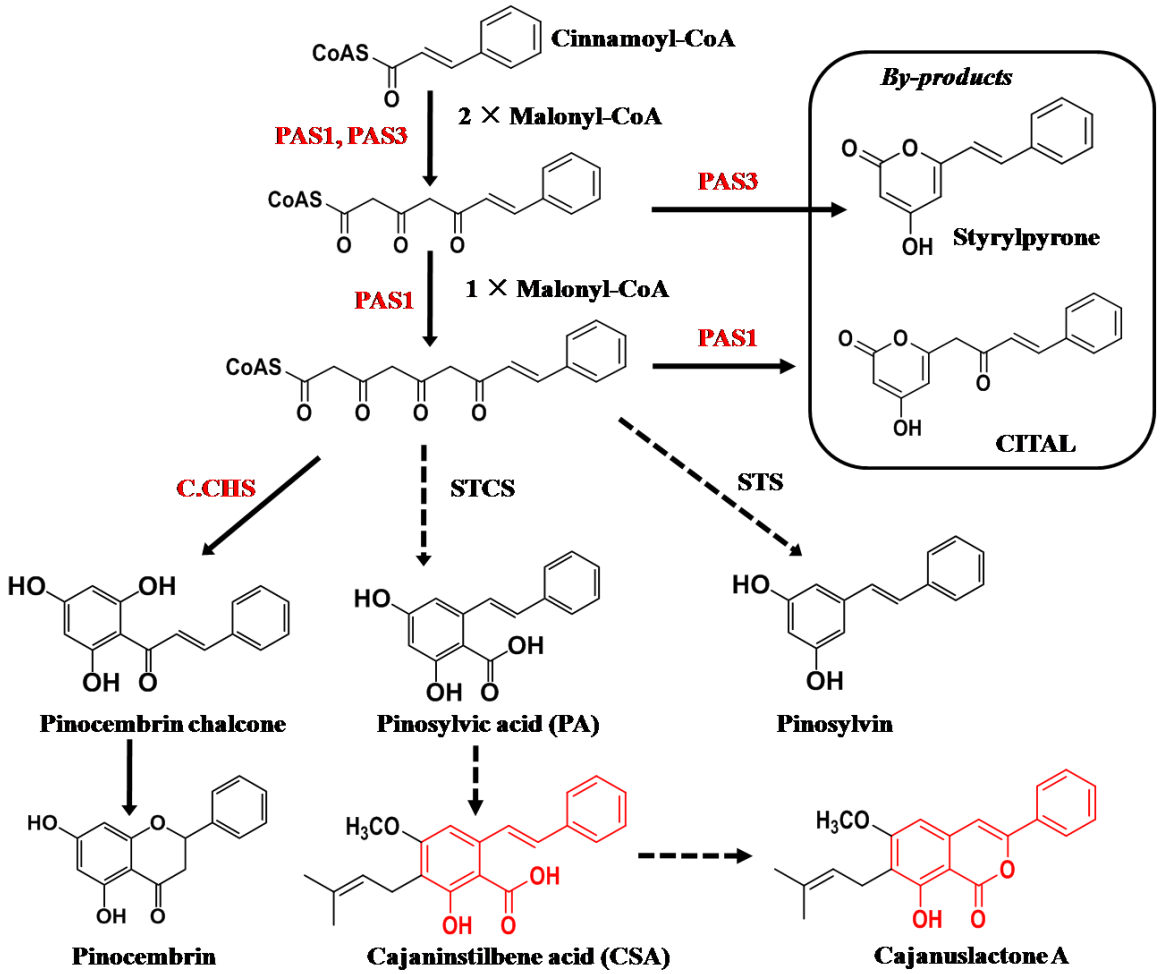


Fig. 61. Proposed cajaninstilbene acid biosynthetic pathway. Solid lines mark confirmed steps, dashed lines so far undetectable steps.

2. Pinosylvic acid synthase (PAS)

For the biochemical characterization of the recombinant PASs, the proteins were used immediately after purification on Ni-NTA agarose. Storage of the enzymes at -20°C was not possible because their activity significantly decreased after freezing, independent of the presence or absence of glycerol. PAS from *C. cajan* shared 64-68% identity with CHS and STSC. However, PAS1 and PAS3 are a rather heterogeneous group and have only 80% identity to each other (Fig. 62). For both PASs, cinnamoyl-CoA is implemented as the preferred starter substrate. PAS1 protein produced CITAL and styrylpyrone as the major products and PAS3 produced only styrylpyrone. These PASs were likely to be responsible for the biosynthesis of related natural products from *C. cajan*, which start from cinnamoyl-CoA as precursor. Beside cinnamoyl-CoA, the two recombinant enzymes also accepted someother starter substrates; however, the relative activities were much lower. The apparent K_m values, as determined from Lineweaver-Burk plots under optimum conditions, were 7.15 and 6.03 µM for cinnamoyl-CoA and 28.56 and 20.16 µM for malonyl-CoA in case of PAS1 and PAS3, respectively. The kinetic parameters determined for the PASs from *C. cajan* correlate with the nature of the

stilbenecarboxylate derivatives in the plant. Both the K_m and K_{cat} values obtained with cinnamoyl-CoA are similar between the two enzymes.

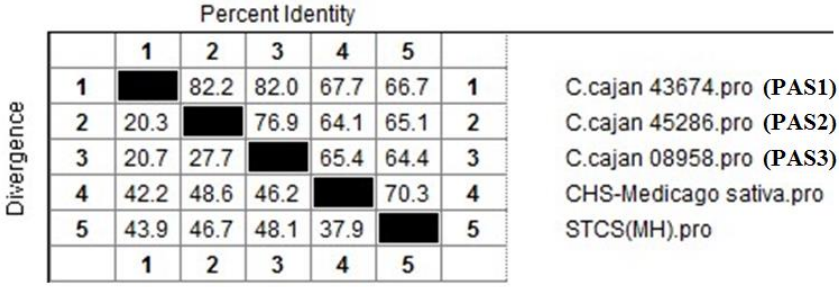


Fig. 62. Alignment of PAS amino acid sequences with CHS (*Medicago sativa*) and STCS (*Hydrangea macrophylla*)

It is interesting to note that CITAL and styrylpyrone, the major products of PASs, were not only detected as the reaction products of CHS of *C. cajan*, but also as those of *Pinus lobata* CHS and *Arachishypogaea* STS when the reaction mixtures were acidified before extraction (Yamaguchi et al., 1999). To date, the formations of these compounds as by-products of CHS or STS have been reported. The apparent reason for past failures to detect these compounds could be the extraction of CHS and/or STS reaction mixtures without acidification. Indeed, when the reaction products of C.CHS, *P. lobata* CHS and *A. hypogaea* STS were extracted without acidification, the detected amount of styrylpyrone was reduced drastically and CITAL could not be detected at all. The chromatograms looked as if naringenin or resveratrol was produced as the sole product. Only one research group (Zuurbier et al., 1993) was aware of the presence of a hydrophilic product in the CHS reaction mixture, which could only be detected when the reaction mixture was acidified before extraction. Although its exact structure was not elucidated and reported, the above circumstantial evidence suggests that this hydrophilic product was CITAL. Judging from the fact that CITAL or its derivatives have been rarely reported from any plant source other than *C. cajan*, it seems that CITAL is a common by-product of all CHSs and STSs acting on cinnamoyl-CoA but its production is limited to *in vitro* reactions. Recently, CHS and STS have been shown to produce 4-hydroxy-2-pyrone derivatives *in vitro* with the nonphysiological substrates isovaleryl-CoA and isobutyryl-CoA. Even though CITAL and styrylpyronewere found as the common by-product of CHS and STS in this study, there are some good reasons to believe that PAS proteins are distinct enzymes which catalyze two or three consecutive decarboxylative condensations of cinnamoyl-CoA with malonyl-CoAs without aromatic ring formation (Akiyama, 1999). The first is that leaves of *C. cajan*, from which the cDNA was cloned, contain a large amount of stilbenecarboxylates and their derivatives (e.g. cajaninstilbene acid), which are probably derived from cinnamoyl-CoA. The second is that the PAS cDNAs were cloned together with C.CHS, the normal CHS superfamily members. The third is that PASs protein produced CITAL and styrylpyrone as the major products *in vitro*. With all this evidence we suggest that the PASs probably provide such a precursor and CITAL is a derailment product in the absence of the subsequent reactions.
In data bases (NCBI) more than 859 nucleotide sequences have been reported for plant PKSs and several PKS crystall structures have been published (Ferrer et al., 1999; Austin

et al., 2004a; Jez et al., 2000a; Morita et al., 2007). This is also true for a bacterial type III PKS (Austin et al., 2004b). There are no significant differences in the conformation of these crystall structures. PKSs commonly form a symmetric dimer displaying an αβαβα five-layered core and each monomer contains an independent active site. Dimerization is required for activity and an allosteric cooperation type between the two active sites from the monomers was suggested (Tropf et al., 1995). Furthermore, it was found that the Met 137 (numbering in *M. sativa* CHS) in each monomer helps to shape the active site cavity of the adjoining subunit (Ferrer et al., 1999).

The basic principle of the reaction mechanism consists of the use of a starter CoA-ester to perform sequential condensation reactions with two carbon units from a decarboxylated extender, usually malonyl-CoA. A linear polyketide intermediate is formed which is folded to form an aromatic ring system (Schröder, 1999). In particular, the active site is composed of a CoA-binding tunnel, a starter substrate-binding pocket and a cyclization pocket (Austin et al., 2004). Three residues conserved in all the known PKSs define this active site: Cys 164, His303 and Asn336. Each active site is buried within the monomer and the substrates enter via a long CoA-binding tunnel. Cys164 is the nucleophile that initiates the reaction and attacks the thioester carbonyl of the starter resulting in transfer of the starter moiety to the cysteine side chain. Asn336 orients the thioester carbonyl of malonyl-CoA near His303, with Phe215 providing a nonpolar environment for the terminal carboxylate that facilitates decarboxylation. Resonance of the enolate ion to the keto form allows for condensation of the acetyl carbanion with the enzyme-bound polyketide intermediate. Phe215 and Phe265 perform as gatekeepers (Austin and Noel, 2003). The recapture of the elongated starter-acetyl-diketide-CoA by Cys164 and the release of CoA set the stage for additional rounds of elongation, resulting in the formation of a final polyketide reaction intermediate. Later an intramolecular cyclization of the polyketide intermediate takes place (Abe, et al., 2003a; Jez et al., 2000b; Jez et al., 2001a). The GFGPG loop is a conserved region on plant PKSs that provides a scaffold for cyclization reactions (Austin and Noel, 2003).

The remarkable functional diversity of the PKSs derives from small modifications in the active site, which greatly influence the selection of the substrate, number of polyketide chain extensions and the mechanism of cyclization reactions. The volume of the active site cavity influences the starter molecule selectivity and limits polyketide length. The 2-PS cavity is one third the size of the CHS cavity. The combination of three amino acid substitutions (Thr197Leu, Gly256Leu and Ser338Ile) in the CHS sequence changed the starter molecule preference from *p*-coumaroyl-CoA to acetyl-CoA and resulted in formation of a triketide instead of a tetraketide product (Jez et al., 2000a). From homology modeling studies, it was found that the cavity volume of octaketide synthase (OKS; Abe et al., 2005b) and aloesone synthase (ALS;Abe et al., 2004a) is slightly larger than that of CHS, while that of pentaketide chromone synthase (PCS) is almost as large as that of ALS (Abe et al., 2005a). The replacing of the residues Ser132Thr, Ala133Ser and Val265Phe fully transformed the ACS to a functional CHS (Lukacin et al., 2001). The change from His166-Gln167 to Gln166-Gln167 converted the STS from *A. hypogaea* to a dihydropinosilvin synthase (Schröder and Schröder, 1992). It was also shown that Gly256, which resides on the surface of the active site, is involved in the chain-length determination of CHS (Jez et al., 2001b). While in ALS, Gly256 determines starter substrate selectivity, Thr197 which is located at the entrance of the buried pocket controls

polyketide chain length. Ser338 in proximity of the catalytic Cys164 guides the linear polyketide intermediate to extend into the pocket, leading to the formation of a heptaketide (Abe et al., 2006b).

The cyclization specificities in the active sites of CHS and STS are due to electronic effects of a water molecule rather than steric factors (Austin et al., 2004a). In BAS, the residue Ser338 is important in the steric guidance of the diketide formation reaction and probably BAS has an alternative pocket to lock the coumaroyl moiety for the diketide formation reaction (Abe et al., 2007). Dana et al. (2006) analyzed mutant alleles of the *Arabidopsis thaliana* CHS locus by molecular modeling and found that changes in the amino acid sequence in regions not located at or near residues that are of known functional significance can affect the architecture, the dynamic movement of the enzyme, and the interactions with others proteins. There may also be dramatic effects on enzyme function.

Although PAS1 and PAS3 catalyze only chain elongation of substrates without aromatic ring formation in a polyketide pathway, they are distinct enzymes which catalyze different reactions. PAS1 produced CITAL, which is the product after three condensations of the substrates, and styrylpyrone, which is the product after two condensations. However, PAS3 produced only styrylpyrone. This means they are different in the process of chain elongation. As described above, the percentage similarity of the amino acid sequences between PAS1 and PAS3 is about 80%, and comparison of their sequences indicated that there are many amino acid residues which show no apparent similarity in charge and size. Among them, some residues which might be responsible for the differences in enzyme function could be pointed out. Some of the different areas might be responsible for controlling chain elongation.

Corresponding active site cavity residues of CHS are shown in Table16 and they were also marked on the sequences of PAS1 and PAS3 (Fig. 63). When comparing the amino acid sequences of the two proteins, four different amino acids between them were Thr 132, Ser 133, Thr194 and Thr197 (numbering in *M. sativa* CHS). These residues are among the essential amino acids of the polyketide elongation pockets of CHSs. Differences of amino acid residues might cause different product formation by two enzymes. A more detailed study viasite-directed mutagenesis of PAS should be conducted in the future and may reveal the origin of product specificity.

Table 16. Active site cavity residues of CHS. (modified from Klundt et al., 2009)

CHS *M. sativa*	
Initiation pocket	**Elongation pocket**
Leucine 214	Threonine 132
Isoleucine 254	Serine 133
Glycine 256	Threonine 194
Phenylalanine 265	Valine 196
Catalytic triad	Threonine 197
Cysteine 164	Glycine 216
Histidine 303	Leucine 263
Asparagine 336	Serine 338

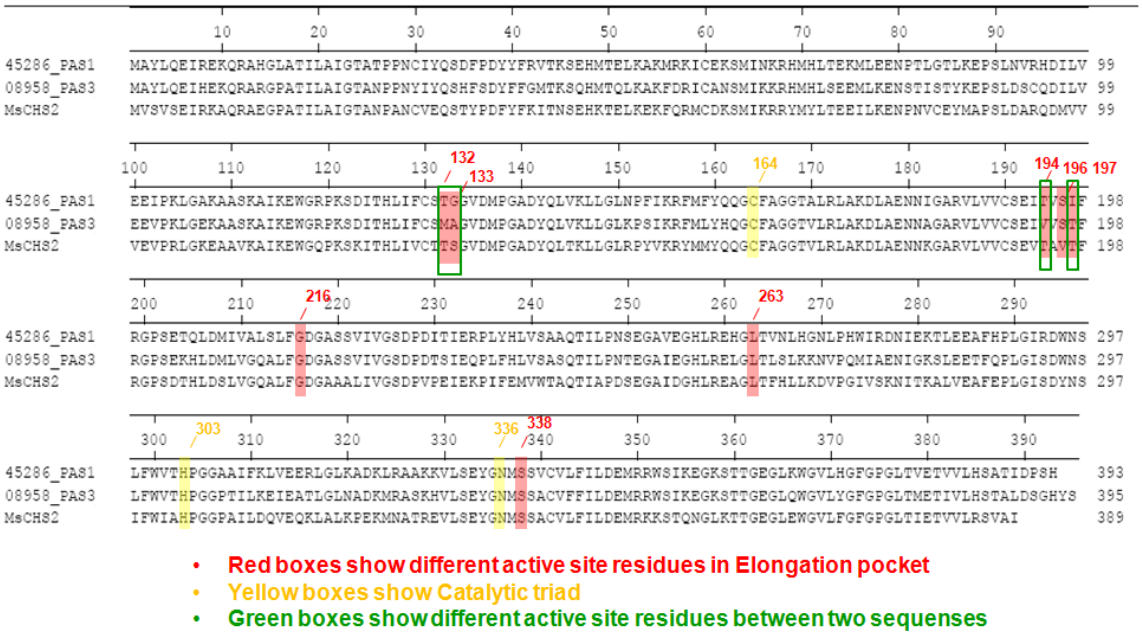


Fig. 63. Alignment of the amino acid sequences of PASs from *C. cajan* with CHS from *Medicago sativa*

3. Phylogenetic characterization of identied PKSs from *C. cajan*

To examine the relatedness among the new cloned genes and some known PKSs, a Maximum Likelihood phylogenetic tree was constructed using homologs of PASs and C.CHS in addition to other PKS homologs.In a phylogenetic tree of type III PKSs (Fig. 64), the amino acid sequences from gymnosperms and angiosperms formed an individual cluster, separated from the PKSs occurring in ferns and mosses (Liu et al., 2003, 2007). Within the angiosperms and gymnosperms, the PKS sequences diverged into two clusters. One group consisted of CHSs includingthe enzymes from different species.CHS is ubiquitously distributed in higher plants and is the prototype enzyme of the type III PKS superfamily (Schröder, 1999; Austin and Noel, 2003). The second cluster was formed by typeIII PKS enzymes, which functionally diverge from CHS. This group included PAS.

The two clusters originate from the duplication of a common ancestral gene prior to the speciation of the angiosperms and gymnosperms (Fig. 64).

Among the functionally diverse type III PKS phenotypes evolved by plants are PAS1 and PAS3 which group together closely in the phylogenetic tree (Fig. 63). PAS1 and PAS3 are subbranches of the same higher branch, indicating that these two cinnamic acid-specific type III PKSs result from a relatively recent functional divergence of a common ancestral gene. In the course of evolution, some enzymes such as stilbene carboxylate synthase (STCS), stilbene synthase (STS) and benzalacetone synthase (BAS) may have evolved several times independently, as indicated by their presence in several clusters in the previously prepared trees (Tropf et al., 1995; Yamazaki et al., 2001). As shown in the present tree, all the C.CHSs were assigned to the CHS group, except for C.CHS2. It is hard to predict why C.CHS2 was found within the functionally diverse group of PKSs and close to HmSTCS and HmCTAS. However, the function of C.CHS2 was clearly confirmed as CHS using the *in vitro* enzymatic incubation. As expected, three C.CHS sequences (C.CHS1, C.CHS3 and C.CHS4) grouped together with their probe, MsCHS2, to form a clearly separated cluster. It was obvious that C.CHS1, C.CHS3 and C.CHS4 are quite distant from C.CHS5 (Fig. 64). They occupy distant branches, which complies with the observation that the *C.cajan* CHSs show diverging substrate selectivities and thus have undergone more independent evolutionary events. Cloning of other CHSs from related species is a perquisite to trace back the ancestor and the exact pylogenetic relationships between the genes. The hypothesis that a gene duplication event occurred and was followed by divergent evolution needs more detailed studies of PKSs in *C. cajan* and corresponding genes in closely related species.

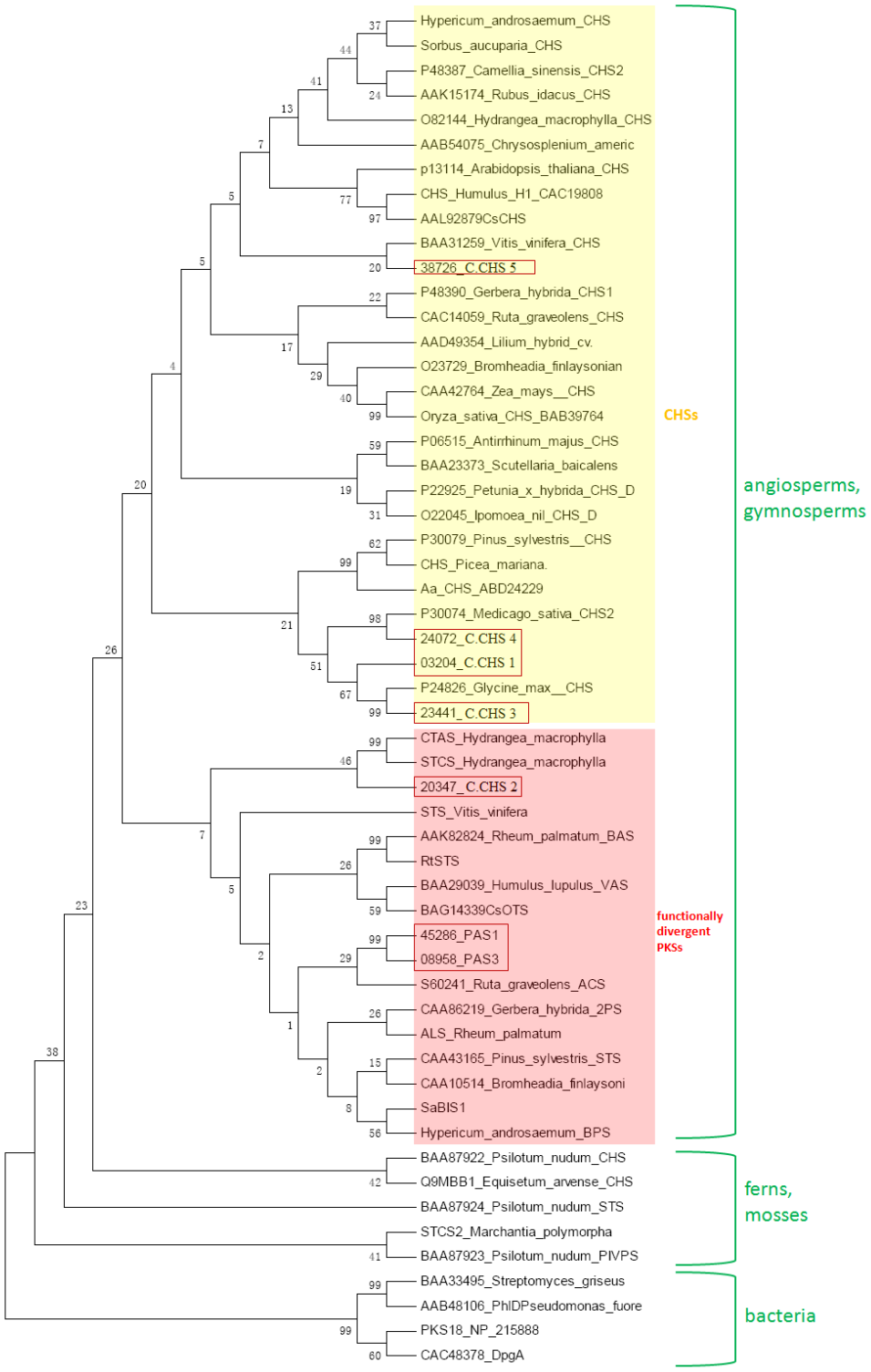


Fig. 64. Neighbour-joining tree of type III PKSs. CHSs (yellow) and non-CHSs (red) from angiosperms and gymnosperms divide into two clusters. Four functionally studied bacterial type III PKSs were used to root the tree. Numbers at the forks are bootstrap values from100 replicates. ACS, acridone synthase; ALS, aloesone synthase; BAS, benzalacetone synthase; BBS, bibenzyl synthase; BIS, biphenyl

synthase; BPS, benzophenone synthase; CHS, chalcone synthase; CTAS, 4-coumaroyl triacetic acid lactone synthase; DpgA, 3,5-dihydroxyphenylacetate synthase; OTS, octaketide synthase; PhID, acetylphloroglucinol synthase; 2-PS, 2-pyrone synthase; STS, stilbene synthase; STCS, stilbene carboxylate synthase; VAS, valerophenone synthase (Liu et al., 2003, 2007).

4. Polyketide cyclase involved in stilbenecarboxylate biosynthesis

As mentioned above, the studies to detectthe STCS activity of PAS in the presence of potential polyketide cyclases failed. However, it may also be possible that a STCS activity of the CHS-related protein was not detected because the *in vitro* incubations lacked some conditions that are necessary for the correct function of the protein. Among these factors, the presence of a specific cyclase which coacts with PAS to produce stilbenecarboxylate is a very reasonable explanation, especially because its functional mechanism has been well demonstrated in *Cannabis sativa* (Gagne et al., 2012). The model of plant polyketide biosynthesis was that carbon scaffolds are assembled exclusively by the activity of type III PKS enzymes (Yu et al., 2008). The discovery of a polyketide cyclase indicates a variant catalytic route inplants, in which cyclase enzymes function cooperatively with typeIII PKSs. It also demonstrates that polyketide cyclases, which until now were only known to partner with type II PKSs inbacterial polyketide pathways, are also found in biosynthetic pathways involving type III PKSs. Although cannabinoid biosynthesis may be unique in its requirement for a cyclase enzyme, similar polyketide cyclases may act in other plants because some possess the hydrophobic cleft and conserved residues. The role of such enzymes may be confined to the formation of polyketide products that retain a carboxylic acid moiety. Examples of plant metabolites formed by C2–C7 aldol condensation with carboxylate retention are the anacardic acids in cashew and gingko and stilbene carboxylates in *Hydrangea* species and liverworts (Kozubek et al., 1999; Gorham et al., 1995). It is worth noting that a rice type III PKS synthesizes long-chain alkylresorcinolic acids without the need for a cyclase enzyme (Matsuzawaet al., 2010). Another indication that some plant polyketide pathways may require cyclases is the formation of α-pyrones when recombinant type III PKS senzymes are assayed *in vitro*, e.g., *Hypericum perforatum* octaketidesynthase (Karppinen et al. 2008). Some possible explanations for different results in our case may be suggested as follows. One is the presence of other polyketide cyclase clones specific for stilbenecarboxylate biosynthesis, which could not yet be obtained by PCRs under the conditions employed in thisstudy. The second is that the identified PASs in our study are not the related PKSs responsible for stilbenecarboxylate biosynthesis in this plant. Polyketide cyclases generally had no intrinsic PKS activity and functioned only in concert with PKSs to help polyketide assembly (Akiyama et al., 1999). The third is that the presence of a specific cyclase, which coacts with the required PKS to produce stilbenecarboxylate, must be associated to distinct incubation conditions. The close association with the cyclase protein and the correct cyclization of the linear polyketide intermediate with carboxylate retention might need the help of yet unknown factors. The *in vitro* incubations so far performed possibly lacked some factors that are necessary for the correct function of the protein. Further investigations are required to uncover the actual mechanism of stilbenecarboxylate biosynthesis in this plant species.

V. Perspectives

Stilbenecarboxylate is a kind of unique secondary metabolite, which is rare in plants. Only a limited number of compounds appear to be known and have been, like stilbenes, isolated from a small but diverse collection of plants.
While candidate type III PKS enzymes have been cloned from some species, *in vitro* reconstitution of stilbenecarboxylate biosynthesis remains to be problematic. The detection of stilbenecarboxylate synthase (STCS) activity and its molecular characterization has not ended the debate about the origin of the carboxyl group of stilbenecarboxylate. However, the existence of stilbenecarboxylates in *C. cajan* provided us a new resource to do further studies on the STCS reaction. Gathering knowledge on the biosynthesis of stilbenecarboxylates has just recently started and more studies are needed before the pathways are fully resolved.

- Unlike protein extracts from leaves, cell-free extracts from *C. cajan* cell cultures might be capable of forming the stilbenecarboxylate *in vitro* after the treatment with elicitors. Thus, the cultivated cells will provide an interesting system to gain deeper insight into stilbenecarboxylate biosynthesis.
- More efforts need to be invested into the detection of stilbenecarboxylate synthase activity with crude protein extract from different plant materials of *C. cajan*. This will help in understanding the biosynthesis of stilbenecarboxylate and other secondary metabolites from *C. cajan* in the future.
- It might also be possible that the presence of some other PKS clones in the genome of *C. cajan* are specific for stilbenecarboxylate biosynthesis, which could not yet be obtained by PCRs under the conditions employed in this study. Further attempts of cloning new type III PKS genes from *C. cajan* seem to be necessary.
- Further attempts have also to be done to clone other polyketide cyclase clones, which might be specific for stilbenecarboxylate biosynthesis in *C. cajan*. The presence of a specific cyclase, which coacts with CTAS to produce stilbenecarboxylate, is a very reasonable explanation.
- Several points suggest that stilbenecarboxylate biosynthesis by CHS-related enzymes may require functional specializations that are not common in other members of the enzyme family. One may be the presence of a reducing reaction because it seems to be common in stilbenecarboxylate biosynthesis. It was shown for the CHS process but is not known for other proteins of the family (Akiyama et al., 1999). Another unique property of the STCS is the extremely tight control of decarboxylation reactions that was not found with other enzymes of the family (Austin et al., 2004). It maybe possible that the stereo-electronic control to avoid the loss of the terminal carboxyl group concomitant with the STS-type ring folding is a critical part of the STCS reaction. It is tempting to speculate that the unusual His in the position next to the active site Cys participates in this control. Mutagenesis studies showed that it is important for the structure and function of the enzyme, but the mechanisms are not clear, and modeling studies based on the CHS structure failed to provide decisive clues. It is noteworthy, however, that a recent publication (Dreier et al., 2001) reported findings that suggest further experiments with the STCS. It described that a

His directly adjacent to the active site Ser of a malonyl-CoA:ACP transacylase could functionally replace the Ser as malonyl acceptor, and the work presented evidence for a malonyl-imidazole adduct. It would be interesting to test whether a similar role in malonyl-binding is possible with the histidine in STCS.

VI. Summary

As a part of the continuous interest in the secondary metabolism of *C. cajan*, we established various *in vitro* cultures of the plant and characterized their chemical composition, aiming at elucidating the biosynthetic pathways of the major metabolites.

- *C. cajan* not only contains stilbenecarboxylate but also other kinds of secondary metabolites with excellent biological activities, such as flavonoids, stilbenes and isocoumarins. Although these constituents were extensively studied on the phytochemical level, their biosynthesis is poorly investigated. The occurrence of cajaninstilbene acid and other stilbenecarboxylate derivatives in *C. cajan* provides a new resource to do further studies on the STCS reaction. The aim of the work was to study the biosynthesis of stilbenecarboxylate at both the enzyme and the gene level. The biosynthesis of cajaninstilbene acid has been proposed to use a polyketide pathway. The first intermediate is pinosylvic acid, which was postulated to be synthesized by a stilbenecarboxylate synthase that uses cinnamoyl-CoA as starter substrate. The postulated CHS and STS reactions are also included in this study.
- Seeds of *C. cajan* were germinated on MS medium and the resulting seedlings were used to establish callus and cell suspension cultures. After optimization of the medium, cell cultures were grown in MS medium supplemented with 0.3 mg/L BA, 0.1 mg/L NAA under a 16 hr photoperiod. Subculturing was performed at 21-day-intervals. Typical growth curves were observed. The long-term maintenance of *in vitro* plants, callus and cell suspension cultures without apparent changes in the growth rates was achieved for further experiments.
- The time courses of accumulation of individual compounds (pinostrobin, longistyline A and cajaninstilbene acid) in *C. cajan* callus and cell suspension cultures as well as *in vitro* leaves revealed that the leaves have the best ability to form cajaninstilbene acid as major constituent (10.45 mg/g dry weight). The compound was absent from callus and cell suspension cultures, while pinostrobin was formed in all three materials. The highest contents of all three compounds were observed during the growth of *in vitro* leaves.
- Variation in content of cajaninstilbene acid during growth of *C. cajan* seedlings revealed that the overall content of this metabolite in leaves was much higher than those in other parts of seedlings. In a comprehensive consideration of the temporal and spatial variations in the content of cajaninstilbene acid, the optimum harvesting time was day 35 after the emergence of leaves.
- Elicitation studies in cell suspension cultures of *C. cajan* were also carried out. All the elicitors (chitosan, methyljasmonate, yeast extract, and sodium chloride) were used with various concentrations and screened for their influence on cajaninstilbene acid production in *C. cajan* cell suspension cultures. None of the elicitors showed any significant increase in metabolites production.
- Using a candidate gene approach, CHS from *Medicago sativa* was selected as a probe to search the *C. cajan* genome sequence for possible candidates participating in stilbenecarboxylate formation. cDNAs encoding seven PKSs were successfully cloned, heterologously expressed, and functionally characterized. The candidates which were

two PASs and five C.CHSs can be subdivided in two groups based on the phylogenetic analysis including the probe CHS and the known STCS from *Hydrangea macrophylla*.

- For PAS enzymes, PAS1 produced CITAL and styrylpyrone in the enzyme assay with cinnamoyl-CoA and malonyl-CoA as the substrates. The cDNA has an ORF of 1182 bp, which encodes 394 amino acids. It shares 67.7 % identity with CHS on the amino acid level. The enzyme has broad substrate specificity. Based on the catalytic efficiency (K_{cat}/K_m), cinnamoyl-CoA was the best substrate. The pH and temperature optima were 7.0 and 35°C, respectively. PAS3 produced only styrylpyrone in the enzyme assay with cinnamoyl-CoA and malonyl-CoA as the substrates. The cDNA has an ORF of 1188 bp, which encodes 396 amino acids. It shares 65.4 % identity with CHS on the amino acid level. The enzyme also has broad substrate specificity. Based on the catalytic efficiency (K_{cat}/K_m), cinnamoyl-CoA was the best substrate. The pH and temperature optima were 7.0 and 35°C, respectively.
- For C.CHSs, detection of pinocembrin together with CTAL and styrylpyrone as products in the enzyme assays with cinnamoyl-CoA and malonyl-CoA as substrates proved that C.CHS1-5 are indeed CHS clones. The cDNAs have ORFs of 1167, 1158, 1167, 1170 and 1176 bp in length for C.CHS1, C.CHS2, C.CHS3, C.CHS4 and C.CHS5, respectively. They share 77-91 % identity with CHS in the database on the amino acid level. Based on the catalytic efficiency (K_{cat}/K_m), C.CHS1, C.CHS2 and C.CHS4 preferred cinnamoyl-CoA as starter substrate and C.CHS3 and C.CHS5 exhibited highest activity with *p*-coumaroyl-CoA.
- The polyketide cyclases that have been mentioned in recent studies (Gagne et al., 2012) were used as probes to search the *C. cajan* genome sequence for possible polyketide cyclases participating in stilbenecarboxylate formation. cDNAs encoding 11 candidates possibly involved were successfully cloned and heterologously expressed. The candidates can be subdivided in three groups based on three probe cyclases (CHI-like, DABB, and Betv1-like), which were five DACs, five BECs and one CHC. Co-incubation of PASs with the recombinant cyclases, however, resulted in no additional products or increases in original product amounts. All the enzyme assays performed failed to detect activity for stilbenecarboxylate biosynthesis.
- Expression of the PAS and C.CHS genes in different plant tissues and in leaves at various growth periods was studied by RT-PCR using gene-specific primers. The distribution of metabolic products correlated with the expression patterns of the PKS transcripts, as determined by expression analysis.

VII. References

Abe I, Abe T, Wanibuchi K and Noguchi H. (2006a). Enzymatic formation of quinolone alkaloids by a plant type III polyketide synthase. *Org Lett.* 8: 6063-6065.

Abe I, Oguro S, Utsumi Y, Sano Y and Noguchi H. (2005b). Engineered biosynthesis of plant polyketides: Chain length control in an octaketide-producing plant type III polyketide synthase. *J Am Chem Soc*. 127: 12709-12716.

Abe I, Sano Y, Takahashi Y and Noguchi H. (2003a). Site-directed mutagenesis of benzalacetone synthase: The role of Phein 215 plant type III polyketide synthases. *J Biol Chem.* 278: 25218-25226.

Abe I, Takahashi Y and Noguchi H. (2002). Enzymatic formation of an unnatural C_6-C_5 aromatic polyketide by plant type III polyketide synthases. *Org Lett.* 4: 3623-3626.

Abe I, Takahashi Y, Lou W and Noguchi H. (2003b). Enzymatic formation of unnatural novel polyketides from alternate starter and nonphysiological extension substrate by chalcone synthase. *Org Lett.*5: 1277-1280.

Abe I, Takahashi Y, Morita H and Noguchi H. (2001). Benzalacetone synthase: A novelpolyketide synthase that plays a crucial role in the biosynthesis of phenylbutanones in Rheum palmatum. *Eur J Biochem.* 268: 3354-3359.

Abe I, Utsumi Y, Oguro S and Noguchi H. (2004a). The first plant type III polyketide syn that catalyzes formation of aromatic heptaketide. *FEBS Lett.*562: 171-176.

Abe I, Utsumi Y, Oguro S, Morita H, Sano Y and Noguchi H. (2005a). A plant type III polyketide synthase that produces pentaketide chromone. *J Am Chem Soc.*127:1362-1363.

Abe I, Watanabe T and Noguchi H. (2004b). Enzymatic formation of long-chain polyketide pyrones by plant type III polyketide synthases. *Phytochemistry.* 65: 2447-2453.

Abe I, Watanabe T and Noguchi H. (2005c). Chalcone synthase superfamily of type III polyketide synthases from rhubarb (Rheum palmatum). *Proc Jpn Acad.* 81(B): 434-440.

Abe I, Watanabe T, Lou W and Noguchi H. (2006b). Active site residues governing substrate selectivity and polyketide chain length in aloesone synthase. *FEBS J.* 273: 208-218.

Abe T, Morita H, Noma H, Kohno T, Noguchi H and Abe I. (2007). Structure function analyses of benzalacetone synthase from Rheum palmatum. *Bioorg Med Chem Lett.* 17: 3161-3166.

Abe T, Noma HA , Noguchi H and Abe I. (2006c). Enzymatic formation of anunnatural methylated triketide by plant type III polyketide synthases. *Tetrahedron Lett*. 47: 8727-8730.

Adawadkar PD and ElSohly MA. (1981). Isolation, purification and antimicrobial activity of anacardic acids from Ginkgo biloba fruits. *Fitoterapia*. 53: 129-135.

Adu-Gyamfi JJ, Fujita K and Ogata S. (1990). Phosphorus fractions in relation to growth in C.cajan (*Cajanus cajan* (L) Millsp.) at various levels of Psupply. *Soil Sci. Plant Nutr.* 36: 531-543.

Ahsan R, Islam M. (2009). Hepatoprotective activity of methanol extract of some medicinal plants against carbon tetrachloride-induced hepatotoxicity in rats. *Euro J Sci Res.* 37:302-310.

Akiyama T, Shibuya M, Liu HM, Ebizuka Y. (1999). *p*-Coumaroyltriaceticacid synthase, a new homologue of chalcone synthase, from Hydrangea macrophylla var. thunbergii. *European Journal ofBiochemistry*. 263: 834-839.

Ames BD et al. (2008). Crystal structure and functional analysis of tetracenomycin ARO/CYC: Implications for cyclization specificity of aromatic polyketides. *Proc Natl Acad Sci USA.* 105:5349–5354.

Anekonda TS. (2006). Resveratrol-a boon for treating Alzheimersdisease. *Brain Res. Rev.* 52: 316-326.

Asakawa Y. (2004). Chemosystematics of the *Hepaticae*. *Phytochemistry*, 65: 623-669.

Asakawa Y. (2007). Biologically active compounds from bryophytes. *Pure Appl. Chem.*, 79: 557-580.

Asahina Y and Asano J. (1930). Uber die constitution von hydrangenol und phyllodulcin (II. Mitteil.). *Chem Ber*. 63: 429-437.

Ashidi JS, Houghton PJ, Hylands PJ, Efferth T. (2010). Ethnobotanical survey and cytotoxicity testing of plants of southwestern Nigeria used to treat cancer, with isolation of cytotoxicconstituents from *Cajanus cajan Millsp.* leaves. *J. Ethnopharmacol.* 128: 501-512.

Athar M, Back JH, Tang X, Kim KH, Kopelovich L, BickersDR and Kim AL. (2007). Resveratrol: a review of preclinical studiesfor human cancer protection. *Toxicol. Appl. Pharmacol.* 224: 274-283.

Austin, MB and Noel JP. (2003). The chalcone synthase superfamily of type III polyketide synthases. *Nat. Prod. Rep.* 20: 79-110.

Austin MB, Bowman ME, Ferrer JL, Schröder J, Noel JP. (2004). An aldol switch discovered in stilbene synthases mediates cyclization specificity of type III polyketide synthases. *Chem. Biol.* 11: 1179-1194.

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA and Struhl K. (1994). Current protocols in molecular biology (Greene Publishing Associates and John Wiley & Sons, Inc. NY).

Baur JA and Sinclair DA. (2006). Therapeutic potential of resveratrol: the in vivo evidence. *Nat. Rev. Drug Discov.* 5: 493-506.

Barry RD. (1964). Isocoumarins development since 1950. *Chem. Rev*. 64: 229-260.

Billek G, Kindl H.(1962). Zur Biosynthese pflanzlicher Stilbene, 2.Mitt.: Die Bildung von Ring A des Hydrangenols. *Monatshefte der Chemie* 93: 814-821.

Bohlmann F and Hoffmann E. (1979). Cannabigerol-ähnliche verbindungen aus Helichrysum umbraculigerum. *Phytochemistry.* 18: 1371-1374.

BorejszaWW and Hrazdina G. (1996). Aromatic polyketide synthases: Purification, haracterization and antibody development to Benzalacetone synthase from raspberry fruits. *Plant Physiol.* 110: 791-799.

Birnboim HC and Doly J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic acids research*. 7: 1513-1523.

Bradshaw HD and Schemske DW. (2003). Allele substitution at a flower colour locus produces a pollinator shift in monkey flowers. *Nature.* 426: 176-178.

Brand S, Holscher D, Schierhorn A, Svatos A, Schröder J and Schneider B. (2006). A type III polyketide synthase from Wachendorfia thyrsiflora and its role in diarylheptanoid andphenylphenalenone biosynthesis. *Planta.* 224: 413-428.

Bradford MM. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry.*72: 248-254.

Burbulis IE, WinkelSB. (1999). Interactions among enzymes of the Arabidopsis flavonoid biosynthetic pathway. *Proc Natl Acad Sci* 96: 12929–12934.

Charles R, Garg SN and Kumar S. (1998). An orsellinic acid glucoside from Syzygium aromatica . *Phytochemistry*. 49: 1375-1376.

Chen H and Chen F. (**2000).** Effects of yeast elicitor on the growth and secondary metabolism of a high-tanshinone-producing line of the Ti transformed *Salvia miltiorrhiza*

cells in suspension culture. *Process Biochemistry*. 35: 837-840.

Chong J, Poutaraud A and Hugueney P. (2009). Metabolism and roles of stilbenes in plants. *Plant Sci.* 117: 143-155.

Chong TM, Abdullah MA, Lai OM, Norini FM, Lajis NH. (2005). Effective elicitation factors in *Morinda elliptica* cell suspension culture. *Process Biochemistry*, 40: 3397-3405.

Cooksey CJ, Dahiya JS, Garratt PJ, Strange RN. (1982). Two novel stilbene-2-carboxylic acid phytoalexins from Cajanus cajan. *Phytochemistry*. 21: 2935-2938.

Cohen SN, Chang AC and Hsu L. (1972). Nonchromosomal antibiotic resistance in bacteria: genetic transformation of Escherichia coli by R-factor DNA. *Proceedings of the National Academy of Sciences of the United States of America*. 69: 2110-2114.

Copeland RA. (2002). Kinetics of Single-Substrate Enzyme Reactions. *In Enzymes (John Wiley & Sons, Inc.)*. 109-145.

Dagert M and Ehrlich SD. (1979). Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. *Gene* 6: 23-28.

Dana CD, Bevan DR and Winkel BSJ. (2006). Molecular modeling of the effects of mutant alleles on chalcone synthase protein structure. *J Mol Model* 12: 905-914.

Dixon RA and Paiva NL. (1995). Stress-lnduced Phenylpropanoid Metabolism. *Plant Cell*.7: 1085-1097.

Duker EG, Jaroszewski JW, Asomaning WA, OppongBF, Brogger C.S. (2004). Antiplasmodial constituentsof *Cajanus cajan*. *Phytother. Res.* 18: 128-130.

Durbin ML, McCaig B and Clegg MT. (2000). Molecular evolution of the chalcone synthase multigene family in the morning glory genome. *Plant Mol Biol* 42: 79-92.

Dreier J, Li Q, Khosla C.(2001). Malonyl-CoA:ACP transacylase from Streptomyces coelicolor has two alternative catalytically active nucleophiles. *Biochemistry* 40: 12407–12411.

Eckermann C, Schröder G, Eckermann S, Strack D, Schmidt E, Schneider B, Schröder J. (2003), Stilbenecarboxylate biosynthesis: a new function in the family of chalcone synthase-relatedproteins. *Phytochemistry*. 62: 271-286.

Ehlting J, Hamberger B, MillionRR, WerckRD. (2006). Cytochromes P450 in phenolic metabolism. *Phytochem. Rev.*5: 239-270.

Ezike AC, Akah PA, Okoli CC, Okpala CB. (2010). Experimental evidence for the antidiabetic activity of *Cajanus cajan* leaves in rats. *J Basic and Clinical Pharm*1:25-30.

Fellermeier M and Zenk MH. (1998). Prenylation of olivetolate by a hemp transferase yields cannabigerolic acid, the precursor of tetrahydrocannabinol. *FEBS Lett*. 427: 283-285.

Fellermeier M, Eisenreich W, Bacher Aand Zenk MH. (2001). Biosynthesis of cannabinoids: incorporation experiments withC-labeled glucoses. *Eur J Biochem*. 268: 1596-1604.

Ferrer JL, Jez JM, Bowman ME, Dixon RA, Noel JP. (1999). Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis, *Nat. Struct. Biol.* 6: 775-784.

Ferrer J, Austin M, Stewart CJ, and Noel J. (2008). Structureand function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiol. Biochem.* 46: 356-370.

Field TS, Lee DW and Holbrook NM. (2001). Why leaves turn red in autumn. The role of anthocyanins in senescing leaves of red-osierdog- wood. *Plant Physiol.* 127: 566-574.

Fischbach MA and Walsh CT (2006). Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem Rev*. 106: 346-349.

Fliegmann J, Schröder G, Schanz S, Britsch L, Schröder J. (1992). Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (Pinus sylvestris), and differential regulation of these and related enzyme activities in stressed plants, *Plant Mol. Biol.* 18: 489-503.

Flores SIJ, Verpoorte R. (2009). Plant polyketide synthases: Afascinating group of enzymes. *Plant Physiol Biochem*., 47: 167-174.

Fu Y, Zu Y, Liu W, Hou C, Chen L, Li S, et al. (2007). Preparative separation ofvitexin and isovitexin from pigeonpea extracts with macroporous resins. *Journal of Chromatography A* 1139: 206-213.

Fujita K, Kai Y, Takayanagi M, El-Shemy H, Adu-Gyamfi JJ and Mohapatra PK. (2004). Genotypic variability of pigeonpea in distribution of photosynthetic carbon at lowphosphorus level**.** *Plant Sci.* 166: 641-649.

Funa N, Awakawa T and Horinouchi S. (2007). Pentaketide resorcylic acid synthesis by type III polyketide synthase from Neurospora crassa. *J Biol Chem.* 282: 14476-14481.

Furuta T, Fukuyama Y, Asakawa Y. (1986). Polygonolide an Isocoumarin from Polygonum-Hydropiper Possessing Anti-Inflammatory Activity. *Phytochemistry*. 25:

517-520.

Gagne SJ, Stouta JM, Liu E, Boubakira Z, Clarka SM and Page JE. (2012). Identification of olivetolic acid cyclase from Cannabissativa reveals a unique catalytic route toplant polyketides. *PNAS.* 31: 12811–12816

Ganeshan S. (2008). Traditional oral care medicinal plants survey of Tamil nadu. *Nat Prod Rad*.7:166-172.

Goodwin PH, Hsiang T and Erickson L. (2000). A comparison ofstilbene and chalcone synthases including a new stilbene synthase gene from Vitis riparia cv. Gloire de Montpellier. *Plant Sci.* 151: 1-8.

Gorham J. (1977). Lunularic acid and related compounds in liverworts, algae and Hydrangea. *Phytochemistry.* 16: 249-253.

Greilhuber J. and Obermayer R. (1998). Genome size variation in Cajanus cajan (Fabaceae): a reconsideration. *Plant Syst. Evol.* 212: 135-141.

Grover JK, Yadav, S, Vats V. (2002). Medicinal plants of India with anti-diabetic potential. *J. Ethnopharmacol*.81: 81-100.

Harborne JB and Williams CA. (2000). Advances in flavonoid research since 1992*Phytochemistry*.55: 481-504.

Heath RJ and Rock CO. (2002). The Claisen condensation in biology. *Nat Prod Rep* 19: 581-596.

Helariutta Y, Kotilainen M, Elomaa P, Kalkkinen N, Bremer K, Teeri T and Albert V.A. (1996). Duplication and functional divergence in the chalcone synthase gene family of Asteraceae: Evolution with substrate change and catalytic simplification. *Proc Natl Acad Sci USA.* 93: 9033-9038.

Holton TA, Brugliera F and Tanaka Y. (1993). Cloning and expression of flavonol synthase from *Petunia hybrida*. *Plant J.* 4: 1003-1010.

Holton TA and Cornish EC. (1995). Genetics and Biochemistry of Anthocyanin Biosynthesis. *Plant Cell*. 7: 1071-1083.

Hopwood DA and Sherman DH. (1990). Molecular genetics of polyketides and its comparison fatty acid biosynthesis. *Annu Rev Genet.* 24: 37-66.

Ibrahim RK, Towers GHN. (1962). Studies of hydrangenol in Hydrangea macrophylla Ser. II. Biosynthesis of hydrangenol from C_{14}-labelled compounds. *Canadian Journal of Biochemistry and Physiology.* 40: 449-457.

Ingham JL, Naturforsch C, (1976). Induced isoflavonoids from fungus-infected stems of pigeon pea (*Cajanus cajan*). *J. Biosci.* 31: 504-508.

Inman, WD, Hoppe, DC. (2002). Compositions containing hypotriglyceridemically active stilbenoids. WO 02/13809.

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A,Choisne, N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F,Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J,Bruye`re C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, CattonaroF, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di GasperoG, Dumas V, Felice N, Paillard S., Juman I., Moroldo M, ScalabrinS, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G,Laucou V, Chatelet P, Merdinoglou D, Delledonne M, Pezzotti M,Lecharny A, Scarpelli C, Artiguenave F, Pe ME, Valle G, MorganteM, Caboche M, Blondon AF, Weissenbach J, Quetier F andWincker P. (2007). The grapevine genome sequence suggests ancestralhexaploidization in major angiosperm phyla. *Nature.* 449: 46-467.

Jeandet P, Paradiso F, Cocina RA, Bolettieri D, Delaunois B,Conreux A, Cordelier S, and Clement C. (2009). Resveratrol inhealth and disease-plant engineering with the stilbene synthase gene. *Cerevisia.* 34: 238-242.

Jeandet P, Douillet AC, Debord S, Sbaghi M, Bessis R andAdrian M. (2002) Phytoalexins from the Vitaceae: biosynthesis, phytoalexin gene expression in transgenic plants, antifungal activity, and metabolism. *J. Agric. Food Chem*. 50: 2731-2741.

Jez JM, Bowman ME, Dixon RA, Noel JP. (2000). Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase. *Nat Struct Biol*.7:786-791.

Kao CM, Katz L and Khosla C. (1994). Engineered biosynthesis of a complete macrolactone in a heterologous host. *Science.* 265: 509-512.

Katsuyama Y, Funa N, Miyahisa I and Horinouchi S. (2007). Synthesis of unnatural flavonoids and stilbenes by exploiting the plant biosynthetic pathway in Escherichia coli. *Chem Biol*. 14: 613-621.

Karppinen K, Hokkanen J, Mattila S, Neubauer P, Hohtola A. (2008). Octaketideproducing type III polyketide synthase from Hypericum perforatum is expressed in dark glands accumulating hypericins. *FEBS J.* 275: 4329–4342.

Ketchum REB, Gibson DM, Croteau RB, Shuler ML. (1998). The kinetics of taxoid accumulation in cell suspension cultures of *Taxus* following elicitation with methyl jasmonate. *Biotechnol Bioeng.* 105: 62-97.

King RE, Bomser JA and Min DB. (2006). Bioactivity of resveratrol. *Compr. Rev. Food Sci. Food Saf.* 5: 65-70.

Klingauf P, Beuerle T, Mellenthin A, El-Moghazy SA, Boubakir Z and Beerhues L. (2005). Biosynthesis of the hyperforin skeleton in Hypericum calycinum cell cultures. *Phytochemistry*. 66: 139-145.

Kodan A, Kuroda H, Sakai F. (2002). A stilbene synthase from Japanese red pine (Pinusdensiflora): implications for phytoalexin accumulation and down-regulation offlavonoid biosynthesis.*Proc. Natl. Acad. Sci.* USA 99: 3335-3339.

Kodan A, Kuroda H and Sakai F. (2001). Simultaneous expression of stilbene synthase genes in Japanese red pine (Pinus densiflora). *J.Wood Sci.* 47, 58-62.

Kozubek A, Tyman JHP (1999). Resorcinolic lipids, the natural non-isoprenoid phenolic amphiphiles and their biological activity. *Chem Rev* 99:1-26.

Kuhn BM, Geisler M, Bigler L, Ringli C. (2011). Flavonols accumulate asymmetrically and affect auxin transport in *Arabidopsis*. *Plant Physiol.*156: 585-595.

KumarK, Thompson JA, Sastry PVSS, Giller KE and Day JM (1987). Measurement of N_2-fixation in field-grown pigeonpea [*Cajanus cajan* (L.) Millsp.] using N_{15}-labelled fertilizer. *Plant Soil.* 101: 107-113.

Langcake P and Pryce RJ. (1977). A new class of phytoalexins from grapevines. *Experientia.* 33: 151-152.

Lans C. (2007). Comparison of plants used for skin and stomach problems in Trinidad and Tobago with Asian Ethnomedicine. *J Ethnobiol Ethnomed.* 33:22-8.

Liswidowati, MF, Hohmann F, Schwer B and Kindl H. (1991). Induction of Stilbene Synthase by Botrytis cinerea in cultured Grapevine cells. *Planta.* 183: 307-314.

Lee JR et al. (2008). Functional characterization of pathogen-responsive protein At-Dabb1 with an antifungal activity from Arabidopsis thaliana. *Biochim Biophys Acta.* 1784:1918-1923.

Liu W, Zu Y, Fu Y, Kong Y, Ma W, Yang M, Li J, Wu N, et al. (2010). Variation in contents of phenolic compounds during growth and post-harveststorage of pigeon pea seedlings. *Food Chemistry.* 121: 732 – 739.

Lu MB, Wong HL, Teng WL. (2001). Effects of elicitation on the production ofsaponin in cell culture of *Panax ginseng. Plant Cell Rep.* 20:674-677.

Lukacin R, Schreiner S and Matern U. (2001). Transformation of acridone synthase to chalcone synthase. *FEBS Lett.* 508: 413-417.

Luo QF, Sun L, Si JY, Chen DH. (2008). Hypocholesterolemic effect of stilbenes containing extract fraction from *Cajanus cajan* on diet induced hypercholesterolemia in mice. *Phytomedicine.* 15:932-939.

Luo M, Liu X, Zu Y, Fu Y, Zhang S, Yao L, et al. (2010). Cajanol, a novel anticancer agent from Pigeonpea [*Cajanus cajan* (L.) Millsp.] roots, induces apoptosis in human breast cancer cells through a ROS mediated mitochondrial pathway. *Chem Biol Interac.* 188:151-160.

MacDonald M.J, Cunha GBD. (2007). A modern view of phenylalanine ammonia lyase, *Biochem. Cell Biol.* 85: 273-282.

Makkar HPS, Siddhuraju P, Becker K. (2007). Methods in molecular biology, Plant secondary metabolites. *Humana press New Jersey* ISBN-10: 1-58829-993-7.

Malikov VM, Yuldashev MP. (2002). Phenolic compounds of plants of the *Scutellaria* genus. Distribution, structure, and properties. *Chem. Nat. Compd.*38: 473-519.

Matsuda H, Shimoda H, Yoshikawa M. (1999). Structure-Requirements of Isocoumarins, Phthalides, and Stilbenes from Hydrangeae Dulcis Folium for Inhibitory Activity on Histamine Release from Rat Peritoneal Mast Cells. *Bioorg. Med. Chem*. 7: 1445-1450.

Matsuzawa M, Katsuyama Y, Funa N, Horinouchi S. (2010). Alkylresorcylic acid synthesis by type III polyketide synthases from rice Oryza sativa. *Phytochemistry* 71: 1059-1067.

Mandel M and Higa A (1992). Calcium-dependent bacteriophage DNA infection. *Biotechnology.*24: 198-201.

Melchior F, Kindl H. (1991). Coordinate and elicitor dependent expression of stilbenesynthase and phenylalanine ammonia-lyase genes in Vitis cv. Optima. *Arch. Biochem. Biophys.* 288: 552-557.

Mol J, Grotewold E and Koes R. (1998). How genes paint flowers and seeds. *Trends Plant Sci.* 3: 212-217.

Moore BS and Hopke JN. (2001). Discovery of a new bacterial polyketide biosynthetic Chempathway. *Biochem*. 2: 35-38.

Morita H, Kondo S, Kato R, Wanibuchi K, Noguchi H, Sugio S, Abe I and Kohno T. (2007). Crystallization and preliminary crystallographic analysis of an acridone-producing novel multifunctional type III polyketide synthase from Huperzia serrata. *Acta Crystallogr*. *F*63: 576-578.

Moss SJ, Martin CJ and Wilkinson B. (2004). Loss of co-linearity by modular polyketide synthases: A mechanism for the evolution of chemical diversity. *Nat Prod Rep.* 21: 575-593.

Napolitano E. (1997). Synthesis of isocoumarins over the last decade: A review. *Org. Prepn. Procedures Intl.* 29: 631-664.

Nicholson RA, David LS, Pan RL, Xin ML. (2010). Pinostrobin from *Cajanus cajan* (L.) Millsp. Inhibits sodium channel-activated depolarization of mouse brain synaptoneurosomes. *Fitoterapia.* 81:826-9.

Nozawa K, Yamada M, Tsuda Y, Kawai K, Nakajima S. (1981). Synthesis of antifungal isocoumarins. II. Synthesis and antifungal activity of 3-substituted isocoumarins. *Chem. Pharm. Bull.* 29: 2491-2495.

Ohwaki Y, Ogino J, Shibano K. (1993). 3-hydroxy-5-methoxystilbene-2-carboxylic acid, a phytotoxic compound isolated frommethanolic extracts of pigeonpea *(Cajanus cajan* Millsp.*)* leaves. *Soil Sci. Plant Nutr.* 39: 55-61.

Oikawa T, Sasaki M. Inose M. Shimamura M, Kuboki H, Hirano S, Kumagai H, Ishizuka M, Takeuchi T. (1997). Effects of cytogenin, a novel microbial product, on embryonic and tumor cell-induced angiogenic responses*in vivo*. *Anticancer Res.*17: 1881-1886.

Okiror MA. 1986. Breeding for Resistance to *Fusarium* wilt. PhD thesis, University of Nairobi

Ono E, FukuchiMM, Nakamura N, Fukui Y, YonekuraSK, Yamaguchi M, Nakayama T, Tanaka T, Kusumi T, Tanaka Y. (2006). Yellow flowers generated by expression of the aurone biosynthetic pathway. *Proc. Natl. Acad. Sci. USA.*103: 11075-11080.

Pal D, Sarkar A, Gain S, Jana S, Mandal S. (2011). CNS depressant activities of *Coccos nucifera* in mice. *Acta Pol Pharm.* 68:249-254.

Park SC et al. (2007). Characterization of a heat-stable protein with antimicrobial activity from Arabidopsis thaliana. *Biochem Biophys Res Commun*. 362: 562–567

Pfeifer BA, Admiraal SJ, Gramajo H, Cane DE and Khosla C. (2001). Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science.* 291**:**1790-1792.

Picerno P, Mencherini T, Lauro MR, Barbato F, and Aquino R. (2003). Phenolicconstituents and antioxidant properties of Xanthosoma violaceum leaves. *Journal of Agricultural and Food Chemistry*. 51: 6423-6428.

Pollastri S and Tattini M. (2011). Flavonols: old compounds for old roles. *Ann. Bot.* 108: 1225–1233.

Preisig MR, Schwekendiek A, Brehm I, Reif HJ and Kindl H. (1999). Characterization of a pine multigene family containing elicitor responsive stilbene synthase genes. *Plant Mol. Biol.* 39: 221-229.

Pryce RJ. (1971). Biosynthesis of lunularic acida dihydrostilbene endogenous growth inhibitor of liverworts. *Phytochemistry*. 10: 2679-2685.

Radauer C, Lackner P, Breiteneder H. (2008). The Betv 1 fold: An ancient, versatilescaffold for binding of large, hydrophobic ligands. *BMC Evol Biol.* 8:286.

Raiber S, Schröder G, Schröder J. (1995). Molecular and enzymatic characterization oftwo stilbene synthases from Eastern white pine (Pinus strobus). A single Arg/Hisdifference determines the activity and the pH dependence of the enzymes. *FEBS Lett.* 361: 299-302.

Rajeev KV, Wenbin C, Yupeng Li, Arvind KB, Rachit KS, et al. (2012). Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nature Biotechnology.* 30: 83-89.

Rao SC, Coleman SW and Mayeux HS. (2002). Forage production and nutritive value of selected pigeonpea ecotypes in the southern great plains. *Crop Sci.* 42: 1259-1263.

Rego TJ and Nageswara RV. (2000). Long-term effects of grain legumes on rainy seasons or ghum productivity in a semi-arid tropical *vertisol. Exp. Agr.* 36: 205-221.

Rix U, Fischer C, Remsing LL and Rohr J. (2002). Modification of post-PKS tailoring the steps rough combinatorial biosynthesis. *Nat Prod Rep*. 19: 542-580.

Rolfs CH, Fritzemeier KH and Kindl H. (1981). Cultured cells of Arachis hypogaea susceptible to induction of stilbene synthase (resveratrol-forming). *Plant Cell Rep.*1:83-85

Saiko P, Szakmary A, Jaeger W and Szekeres T. (2008). Resveratrol and its analogs: defense against cancer, coronary disease and neurodegenerative maladies or just a fad? *Mutat. Res.* 658: 68-94.

Saslowsky, DE, Warek U, Winkel BSJ. (2005). Nuclear localization of flavonoid enzymes in *Arabidopsis*. *J. Biol. Chem.,* 280: 23735-23740.

Sato H, Konoma K, Sakamura S.(1981). Three new phytotoxins produced by Pyrenochaeta terrestris: pyrenochaetic acids A, B and C. *Agric. Biol. Chem*. 45: 1675-1679.

Saxena KB, Kumar RV and Rao PV. (2002). Pigeonpea nutrition and its improvement Quality Improvement in Field Crops. *Food Products Press.* 227-260.

Seshadri TR, (1957). *Current Sci.* (India). *Chem. Abstr.* 52: 857.

Schöppner A, Kindl H. (1984). Purification and properties of a stilbene synthase frominduced cell suspension cultures of peanut. *J. Biol. Chem.* 259: 6806-6811.

Shang XF, He XR, He X, Li MX, Zhang RX, Fan PC, Zhang QL, Jia ZP. (2010). Thegenus *Scutellaria* an ethnopharmacological and phytochemical review. *J.Ethnopharmacol.*128: 279-313.

Shibuya M, Nishioka M, Sankawa U and Ebizuka Y. (2002). Incorporation of three deuterium atoms excludes intermediacy of stilbenecarboxylate in stilbene synthase reaction. *Tetrahedron Lett.* 43: 5071-5074.

Schmidlin L, Poutaraud A, Claudel P, Mestre P, Prado E, Santos-Rosa M, Wiedemann MS, Karst F, Merdinoglu D and Hugueney P. (2008). A stress-inducible resveratrol *O*-methyltransferase involved in the biosynthesis of pterostilbene in grapevine. *Plant Physiology.* 148: 1630-1639.

Siemann EH, Creasy LL. (1992). Concentration of the phytoalexin resveratrol in wine. *Am. J. Enol. Vitic.* 43: 49-52.

Singh S, Mehta A, John J, Mehta P. (2010). Anthelmintic potential of *Andrographis paniculata*, *Cajanus cajan* and *Silybum marianum*. *Pharmacog J.* 2: 71-3.

Stafford HA. (1991). Flavonoid evolution: an enzymic approach. *Plant Physiol.* 96: 680-685.

Stafford HA. (1974). Possible multienzyme complexes regulating the formation of C_6-C_3 phenolic compounds and lignins in higher plants. *Recent Adv. Phytochem.*, 8: 53-79.

Staunton J and Weissman KJ. (2001). Polyketide biosynthesis: A millennium review. *Nat Prod S Rep.* 18: 380-416.

Thompson TB, Katayama K, Watanabe K, Hutchinson CR, Rayment I. (2004). Structural and functional analysis of tetracenomycin F2 cyclase from Streptomyces glaucescens. A type II polyketide cyclase. *J Biol Chem.* 279: 37956–37963.

Tanassov I, Russinova E, Antonov L and Atanassov A. (1998). Expression of an anther-specifics.A chalcone synthase-like gene is correlated with uninucleate microspore development in Nicotiana sylvestri. *Plant Mol Biol.* 38: 1169-1178.

Tian L, Wan SB, Pan QH, Zheng YJ, Huang WD. (2008). A novel plastid localization of chalcone synthase in developing grape berry. *Plant Sci.*, 175: 431-436.

Toda K, Kuroiwa H, Senthil K, Shimada N, Aoki T, Ayabe S, Shimada S, Sakuta M, Miyazaki Y, Takahashi R. (2012). The soybean F3'H protein is localized to the tonoplast in the seedcoat hilum. *Planta*. 236: 79-89.

Thu TT, Mai TTX, Dewaele E, Farsi S., Tadesse Y., Angenon G. and Jacobs M. (2003). *In vitro* regeneration and transformation of pigeonpea [*Cajanus cajan* (L.) Millsp.]. *Mol. Breeding*. 11: 159-168.

Tropf S, Kärcher B, Schröder, G, and Schröder, J. (1995). Reaction mechanisms of homodimeric plant polyketide synthases (stilbene andchalcone synthase): a single active site for the condensing reaction is sufficient for synthesis of stilbenes, chalcones, and 60-deoxychalcones. *J. Biol. Chem.* 270: 7922-7928.

Tropf S, Lanz T, Rensing SA, Schröder J, Schröder G. (1994). Evidence that stilbenesynthases have developed from chalcone synthases several times in the course of evolution, *J. Mol. Evol.* 38: 610-618.

Upadhyay B, Parveen, Dhaker AK, Kumar A. (2010). Ethnomedicinal and ethnopharmaco statistical studies of Eastern Rajasthan. *J Ethnopharmacol.* 129: 64-86.

Valio IFM and Schwabe WW. (1970). Growth and dormacy in Lunularia cruciata (L.) Dum. VIII: The isolation and bioassay of lunularic acid. *J Exp Bot*. 21: 138-150.

Varanda EA, Varella SD, Rampazo RA, Kitagawa RR, Raddi MSG, Vilegas W, dos Santos LC. (2006). Mutagenic and cytotoxic effect of planifolin: A naphthopyranone dimer isolated from *Paepalanthus planifolius*. *Toxicologyin Vitro.*20: 664-668.

Wanibuchi K, Zhang P, Abe T, Morita H, Kohno T, Chen G, Noguchi H and Abe I. (2007). An acridone-producing novel multifunctional type III polyketide synthase from Huperzia serrata. *FEBS J*. 274: 1073-1082.

Wang WX, Pelah D, Alergand T, Shoseyov O, Altman A (2002). Characterization of SP1, a stress-responsive, boiling-soluble, homo-oligomeric protein from aspen. *Plant Physiol.* 130: 865-875.

Wang HL, Wang W, Zhang P, Pan QH, Zhan JC, Huang WD. (2010). Gene transcript accumulation, tissue and subcellular localization of anthocyanidin synthase (ANS) in developing grape berries. *Plant Sci.*179: 103-113.

Wang HL, Wang W, Li H, Zhang P, Zhan JC, Huang WD. (2011). Expression and tissue and subcellular localization of anthocyanidin synthase (ANS) in grapevine. *Protoplasma*. 248: 267-279.

Weisshaar B and Jenkins G I. (1998). Phenylpropanoid biosynthesis and its regulation.

*Curr. Opin. Plant Biol.*1: 251-257.

Whyte AC, Glober JB, Scott JA, Mallock D. (1996). Cercophorins A-C: Novel antifungal and cytotoxic metabolites from the coprophilous funguscercophora areolata. *J. Nat. Prod*. 59: 765-769.

Wiese W, Vornam B, Krause E and Kindl H. (1994). Structural organization and differential expression of three stilbene synthasegenes located on a 13 kb grapevine DNA fragment. *Plant Mol. Biol.*26: 667-677.

Wijeratne EMK, Paranagama PA, Gunatilaka AAL. (2006). Five new isocoumarins from Sonoran desert plant-associated fungal strains *Paraphaeosphaeria quadriseptata* and *Chaetomium chiversii. Tetrahedron.* 62: 8439-8446.

Winkel BSJ. (2004). Metabolic channeling in plants. *Annu. Rev. Plant Biol. 55*: 85-107.

Winkel SB. (2002). Biosynthesis of flavonoids and effects of stress. *Curr.Opin.Plant Biol.* 5: 218–223.

Winkel SB. (2001). Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol.*126: 485-493.

Wu N, Fu K, Fu YJ, Zu YG, Chang FR, Chen YH, et al. (2009). Antioxidant activities of extracts and main components of pigeon pea leaves. *Molecules* 14:1032–1043.

Wu S, Leary SJB, Gleddie S, Eudes F, Laroche A and Robert LS. (2008). A chalcone synthase-like gene is highly expressed in the tapetum of both wheat (Triticum aestivum L.) and striticale (x Triticosecale Wiimack). *Plant Cell Rep.* doi: 10.1007/s00299-008-0572-3.

Wu S, Chappell J. (2008). Metabolic engineering of natural products in plants; tools of the trade and challenges for the future. *Curr Opin Biotechnol.* 19: 145-152

Yoshikawa H, Ichiki Y, Sakakibara K, Tamura H, Suiko M. (2002). The biological and structuralsimilarity between lunularic acid and abscisic acid. *Biosci. Biotechnol. Biochem.* 66: 840-846.

Young ND, Mudge J. and Ellis TN. (2003). Legume genomes: more than peas in a pod. *Curr. Opin. Plant Biol.* 6: 199-204.

Yu O and Jez M. (2008). Nature's assembly line: biosynthesis of simple phenylpropanoids and polyketides. *Plant J.* 54: 750-762.

Yu CK, Springob K, Schmidt J, Nicholson RL, Chu IK, Yip WK, Lo C. (2005). A stilbenesynthase gene (SbSTS1) is involved in host and nonhost defense responses in Sorghum. *Plant Physiol.* 138: 393-401.

Zhang H, Matsuda H, Kumahara A, Ito Y, Nakamura S, Yoshikawa M. (2007). New type of anti-diabetic compounds from the processed leaves of Hydrangea macrophylla var. thunbergii (Hydrangeae Dulcis Folium). *Bioorg. Med. Chem. Lett.*17: 4972-4976.

Zhang DY, Zhang S, Zu YG, Fu YJ, Kong Y, Gao Y, et al. (2010). Negative pressure cavitation extraction and antioxidant activity of genistein and genistin from the roots of pigeon pea. *Separation and Purification Tech.*74: 261-270.

Zhang HL, Matsuda H, Yamashita C, Nakamura S, Yoshikawa M. (2009). Hydrangeic acid from the processed leaves of *Hydrangea macrophylla* var. *thunbergii* as a new type of anti-diabeticcompound. *Eur. J. Pharmacol* .606: 255-261.

Zhang W, Krohn K, Draeger S, Schulz B. (2008). Bioactive Isocoumarins Isolatedfrom the Endophytic Fungus *Microdochium bolleyi J. Nat. Prod.* 71: 1078-1081.

Zheng H, Saxena KB, Chao H, Jianyun Z, Yong G, Xuxiao Z and ShiyingY. (2001). Pigeonpea: An excellent host for Lac production. *ICPN* 8: 58-60.

Zhou K, Wu B, Zhuang Y, Ding L, Liu Z, Qiu F. (2009). Chemical constituents of fresh celery.*Zhongguo Zhong Yao Za Zhi*, 34: 1512-1515.

Zuurbier KWM, Fung SY, Scheffer JJC and Verpoorte, R. (1993). Assay of chalcone synthase activity by high-performance liquid chromatography. *Phytochemistry*. 34: 1225-1229.

VIII. Appendix

1. Sequences

1.1 PAS

PAS1 (C.cajan_45286)
ATGGCATATTTGCAAGAAATACGTGAGAAACAAAGAGCTCATGGCCTTGC
CACCATATTAGCCATTGGTACTGCAACTCCACCCAACTGCATTTACCAAT
CTGACTTTCCCGATTATTATTTCCGAGTCACCAAGAGCGAGCACATGACT
GAGCTCAAGGCCAAGATGAGGAAAATATGTGAGAAGTCGATGATAAATAA
ACGTCACATGCATTTGACAGAAAAAATGCTAGAAGAAAACCCCACCCTTG
GCACTTTAAAGGAACCATCCCTAAATGTTCGCCACGACATACTAGTTGAG
GAGATTCCCAAGCTTGGTGCAAAGGCAGCGTCAAAAGCCATCAAGGAATG
GGGTAGACCTAAGTCAGATATCACTCATCTCATATTTTGCTCCACCGGAG
GTGTAGACATGCCTGGTGCCGACTACCAACTTGTGAAGCTCTTAGGCCTC
AATCCATTCATCAAAAGGTTCATGTTTTACCAACAAGGTTGCTTTGCTGG
TGGCACTGCCCTTCGTCTCGCTAAAGACCTTGCGGAGAACAATATTGGAG
CACGTGTTCTCGTCGTATGTTCTGAGATTACAGTATCCATATTTCGTGGA
CCCTCCGAAACACAGTTGGACATGATAGTTGCACTATCTCTCTTTGGTGA
TGGTGCTTCATCCGTGATCGTTGGATCTGACCCTGACATAACCATTGAAC
GACCACTATATCACCTTGTCTCAGCGGCACAGACAATTCTACCCAACTCT
GAGGGTGCAGTTGAGGGACACTTGCGTGAGCATGGACTCACCGTCAATCT
TCACGGAAATCTTCCTCATTGGATTCGAGACAACATAGAGAAAACCCTTG
AAGAAGCTTTTCACCCACTTGGGATCAGAGACTGGAACTCGTTGTTCTGG
GTAACACACCCTGGTGGCGCTGCAATCTTTAAGCTGGTAGAGGAAAGACT
TGGGTTGAAGGCTGATAAACTGAGAGCAGCTAAAAAAGTTCTAAGTGAGT
ACGGGAACATGTCGAGTGTATGTGTCCTATTTATATTGGATGAGATGAGA
AGGTGGTCTATCAAGGAAGGGAAATCCACCACTGGTGAAGGACTGAAGTG
GGGTGTTTTACATGGGTTTGGCCCTGGCTTGACAGTGGAGACCGTTGTGT
TGCATAGTGCCACCATCGACCCAAGCCATTAG

PAS3 (C.cajan_08958)
ATGGCATACTTGCAGGAAATACATGAGAAACAAAGAGCTCGTGGTCCAGC
CACCATACTAGCCATTGGTACCGCAAATCCACCCAACTACATTTACCAAT
CTCACTTTTCCGATTACTTCTTCGGAATGACCAAGAGCCAACACATGACT
CAACTTAAGGCCAAGTTCGATCGCATATGTGCGAATTCGATGATAAAGAA
ACGTCACATGCACCTGAGCGAAGAAATGTTGAAAGAAAACTCCACTATAA
GCACTTATAAGGAACCATCCTTGGATTCTTGCCAAGACATACTAGTTGAG
GAGGTTCCAAAGCTTGGCGAGAAGGCAGCATCAAAAGCCATCAAGGAATG
GGGCAGACCCAAGTCAGATATCACTCACCTCATATTTTGCTCCATGGCAG
GTGTAGACATGCCTGGTGCGGATTATCAACTTGTGAAGCTCTTAGGTCTC
AAGCCATCCATCAAAAGGTTCATGTTATACCATCAAGGTTGCTTTGCTGG
TGGAACTGTGCTTCGTCTCGCTAAAGACCTTGCTGAGAACAATGCTGGAG
CACGCGTTCTGGTGGTATGTTCTGAGATTGTAGTCAGCACATTTCGTGGA

CCCTCAGAAAAACACTTGGACATGTTGGTGGGACAAGCTCTCTTTGGTGA
TGGTGCCTCATCCGTGATCGTTGGATCCGACCCTGACACGTCCATTGAGC
AACCACTGTTTCACCTTGTTTCGGCATCACAAACAATTCTACCCAACACT
GAGGGTGCAATTGAGGGACACTTGCGAGAACTGGGACTCACATTAAGTCT
CAAGAAAAATGTTCCTCAGATGATCGCAGAGAACATAGGAAAAAGCCTCG
AAGAAACCTTTCAACCACTTGGAATTAGTGATTGGAACTCGTTGTTTTGG
GTAACCCACCCCGGTGGCCCTACAATCTTGAAGGAAATAGAGGCAACACT
TGGGTTAAACGCTGATAAAATGAGAGCATCTAAACATGTCCTAAGTGAGT
ACGGAAACATGTCAAGTGCTTGTGTCTTCTTTATACTAGACGAGATGAGA
AGGTGGTCCATTAAGGAAGGAAAATCCACCACTGGTGAAGGACTACAGTG
GGGTGTTTTATATGGGTTTGGCCCTGGATTGACCATGGAGACCATTGTTT
TGCATAGCACTGCCTTAGACTCAGGCCACTATAGCTAG

1.2 C.CHS

C.CHS1 (C.cajan_03204)
ATGGTGTCTGTTGATGAGATCCGCAAGGCGCAAAGAGCAGAAGGCCCTGC
CACCGTTCTGGCTATCGGCACCGCAACTCCTCCAAACTGTGTGGACCAAA
GTACCTATCCTGACTACTATTTCCGCATCACCAACAGTGAACACAAGACC
GAACTCAAAGAAAAATTCAAGCGCATGTGTGAAAAGTCAATGATTAAGAA
GAGATACATGCACTTGACAGAAGAGATTCTGAAGGAGAATCCAAGCATAT
GCGAGTACATGGCACCTTCATTGGATGCAAGGCAAGACATGGTGGTTGTT
GAAGTGCCAAAGCTGGGGAAAGAGGCTGCAACAAAGGCAATCAAGGAATG
GGGTCAACCTAAATCTAAGATTACTCATCTCATCTTTTGCACCACAAGTG
GTGTGGACATGCCAGGTGCTGACTACCAACTCACCAAGTTGTTGGGTCTT
AGGCCATATGTGAAGCGTTACATGATGTACCAACAAGGTTGTTTCGCTGG
TGGCACAGTTCTACGTTTGGCAAAAGACTTGGCTGAGAACAACAAGGGTG
CACGTGTGTTGGTGGTTTGTTCTGAGATCACTGCAGTCACTTTCCGTGGC
CCAACAGACACTCATCTTGACAGTCTTGTGGGTCAAGCACTGTTTGGAGA
TGGTGCAGCTGCTGTTATTGTTGGTTCAGACCCTTTGCCTGTTGAGAAGC
CTTTGTTTGAATTGGTGTGGACTGCACAAACAATCCTCCCAGACAGTGAA
GGAGCCATTGATGGCCACCTTCGTGAAGTTGGACTCACTTTTCATCTTCT
CAAAGATGTTCCTGGGCTCATCTCAAAGAACATTGAGAAGGCTTTGGTTG
AAGCCTTTCAACCTTTGAACATCTCTGACTACAATTCTATCTTTTGGATT
GCACACCCTGGTGGACCTGCAATTCTGGACCAAGTAGAGGCCAAGTTAGG
CTTAAAGCCAGAGAAGATGAGAGCCACCAGACAGGTGCTAAGCGACTATG
GAAACATGTCAAGTGCGTGCGTGTTGTTTATAATGGACGAGATGAGGAGG
AAGTCAAAGGAAGATGGAAAAGCCACAACAGGAGAGGGTCTTGATTGGGG
TGTGCTCTTTGGTTTTGGACCAGGACTCACTGTTGAGACCGTTGTTCTCC
ATAGTGTTGCAGCTTGA

C.CHS2 (C.cajan_20347)
ATGGCAACGGTCGATGAGATTCACAGTGCACAAAGGGCAAAAGGCCCAGC
AAGTATCTTGGCAATAGGCATTGCCACTCCTCAAAATTGTGTGGAACAAA
GTACCTATCCTGACTACTATTTTCGTATCACAAATAGTGACCACAAGATC
GAACTTAAAAATAAATTCAAACGCATGTGCGAAAATTCAATGATTAAGAA

GCGATACATGCACTTGACAGAAGAGATCCTAAAGAAGAATCCCAATATTG
GTGAGCACATGGCACCTTCATTGGATGCAAGGCAAGAGTTGGTGGTTATG
GAAGTACCAAAGTTAGGAAAGGAGGCAGCAATGAAGGCCATTAAGGAATG
GGGTGAACCTAAGTCTAAAATCACCCACCTCATATTTTGCACCACAAGTG
GTGTGGACATGCCTGGTGCTGACTACCAACTCACAAAGCTCTTAGGCCTT
TGCCCAAATGTGAAACGTTATATGATATACCAACAAGGTTGCTTTGCTGG
TGTTATGGTACTTCGTTTGGCTAAAGATTTGGCTGAAAATAATAAAGGTG
CTCGTGTGTTGATAGTTTGCTCAGAAAACACCATATTCACTTTTCGTGGC
CCAAGTGACACTCATGTTGATAACCTTGTGGGGCAATCACTATTTGGAGA
TGGTGCAACAGCTGTCATTATTGGTTCAGACATAATACCAAAGGTTGAGA
AGCCTTTGTTTCAATTAGTGTGGACTTCACAAACAATCATCCCTGATAGT
GAAGGAGCCTTGGAATTGCATCTTCGTGAAGCTGGGTTGACATTCCATCT
TTTGAAGGAAGTTCCAAATCTTATTTCAAACAACATTGAAAATGCTTTGT
GTGAAGCCTTTCAACCATTGAACATCTCTGATTATAACTCCATATTTTGG
ATAGTACATCCTGGTGGGCGAGCAATTCTAGACCAAATTGAGGCTAAGTT
AGGTTTAAAGGCAGAAAAAATGCGAGCCAGTCGACACATCCTTAGTGAGT
ATGGTAATATGGCAAGTGTTTGTGTGTTTTTCATCATGGATGAAATGAGG
AGGAAGTCAAAAGAAAATGGACTTTCCACAACAGGTTGGGGTGTACTATT
TGGCTTTGGACCTGGAATCACTTTAGAGACTATTGTTCTTCATAGTGTAG
CACTCTAA

C.CHS3

(C.cajan_23441)ATGGTGAGTGTTGAAGATATCCGAAAGGCACAACGTGCAGAA
GGCCCTGC
CACTGTCATGGCTATTGGCACCGCCACTCCTCCCAACTGTGTCGATCAGA
GTACATATCCTGATTATTATTTCCGCATCACAAACAGTGAGCACAAGACT
GAGCTCAAAGAAAAGTTTAAGCGAATGTGTGATAAGTCAATGATCAAGAA
GCGGTACATGTACTTGAATGAAGAGATCCTGAAGGAGAATCCGAGTGTGT
GTGAATACATGGCACCTTCATTGGATGCAAGGCAAGACATGGTGGTTGTG
GAGGTACCAAAGTTGGGAAAAGAGGCTGCAACTAAGGCAATCAAGGAATG
GGGTCAACCCAAGTCAAAGATCACTCATCTCATCTTTTGCACCACCAGTG
GTGTCGACATGCCTGGTGCTGACTATCAGCTCACTAAACTCCTAGGCCTT
CGTCCCTCCGTGAAACGTTACATGATGTACCAACAAGGTTGCTTTGCCGG
TGGCACGGTCCTTCGTTTGGCCAAAGACTTGGCTGAAAACAACAAGGGTG
CTCGTGTGCTTGTGGTTTGCTCCGAGATCACTGCAGTCACATTCCGTGGC
CCAAGTGACACCCATCTTGATAGCCTTGTGGGGCAAGCCCTGTTTGGAGA
TGGTGCTGCTGCAGTCATTGTTGGATCAGACCCTTTGCCAGTTGAAAAGC
CTTTCTTTGAGCTTGTGTGGACTGCACAAACAATCCTTCCGGATAGTGAA
GGGGCTATTGATGGTCACCTTCGTGAAGTTGGACTCACATTCCATCTCCT
CAAGGATGTTCCTGGACTCATCTCTAAGAACATCGAGAAGGCCTTAGTTG
AAGCCTTCCAACCCTTGGGAATCTCTGATTACAACTCTATCTTTTGGATT
GCACACCCTGGTGGCCCTGCAATTTTAGACCAGGTTGAGGCTAAGTTAGG
CTTGAAGCCTGAAAAAATGGAAGCTACTAGACATGTGCTGAGTGAGTATG
GTAACATGTCAAGTGCATGTGTTCTATTCATCTTGGATCAAATGAGGAAG
AAGTCGATAGAAAATGGACTCGGCACAACTGGTGAAGGCCTTGAATGGGG
TGTGCTATTTGGTTTTGGCCCTGGACTCACTGTTGAGACTGTTGTGCTCC

GCAGTGTCACTGTCTAA

C.CHS4 (C.cajan_24072)
ATGGCCAGCGTTGCTGAGATCCGCAAGGCCCAAAGGGCAGAAGGCCCAGC
AACTATCTTGGCTATTGGAACTGCCAATCCACCAAACTGTGTTGACCAGA
GCACCTATCCTGATTTCTACTTCAGAATCACCAACAGCCAGCACATGACT
GAACTCAAAGAGAAATTTCAGCGCATGTGCGACAAGTCTATGATCAAGAG
GAGATACATGTACTTAAATGAAGAGATCTTGAAGGAGAATCCCAACATGT
GCGCTTACATGGCACCATCTTTGGATGCTAGGCAAGACATGGTGGTGGTG
GAGGTACCAAAGCTAGGGAAAGAGGCAGCAGCAAAGGCTATAAAGGAGTG
GGGGCAGCCAAAATCAAAGATCACCCACTTAATCTTCTGCACCACAAGTG
GTGTGGACATGCCTGGTGCTGACTACCAACTCACCAAACTCTTGGGACTT
CGCCCCTATGTGAAGAGGTACATGATGTACCAACAAGGGTGCTTTGCTGG
TGGCACGGTGCTTCGTTTGGCCAAGGATTTGGCTGAGAACAACAAGGGTG
CACGTGTGCTTGTTGTGTGTTCTGAGATCACTGCTGTCACATTCCGTGGC
CCAAGTGACACTCACCTAGATAGTCTTGTGGGGCAAGCACTGTTTGGAGA
TGGTGCGGCTGCAGTGATTGTTGGTTCTGACCCAATTCCACAGGTTGAGA
AGCCTTTGTATGAGCTTGTTTGGACTGCACAAACTATTGCTCCAGACAGT
GATGGTGCTATTGATGGCCACCTTCGTGAAGTTGGCCTCACGTTTCACCT
CCTCAAAGACGTTCCCGGGATTGTCTCTAAGAACATTGACAAGGCACTTT
TTGAGGCCTTCAACCCTTTGAACATCTCTGATTACAACTCCATTTTCTGG
ATTGCACACCCCGGTGGACCAGCAATTCTTGACCAAGTTGAGCAAAAGTT
GGCTCTCAAGCCTGAAAAGATGAGGGCCACAAGGGAAGTGCTCAGTGAGT
ATGGTAACATGTCAAGTGCTTGTGTGCTTTTCATCTTGGATGAGATGAGG
AGGAAATCAGCAAAAGATGGGCTTAAAACCACAGGTGAAGGACTAGAGTG
GGGTGTGTTGTTTGGTTTTGGACCTGGTCTTACTATTGAAACTGTTGTTC
TGCACAGTGTGGCCATATAA

C.CHS5 (C.cajan_38726)
ATGGTGACGGTGGAGGAAATCCGCAGCGCCCAGCGCGCTCATGGCCCCGC
CACCATCTTGGCCTTGGGCACCGCCACTCCCTCCAACTGCATCTCCCAAG
CCGATTACCCCGACTACTACTTCCGCATCACTAACAGCCAACACATGACC
GACCTCAAAGAAAAGTTCAAACGCATGTGTGAAAAGTCGATGATAAAGAA
GCGATACATGCACCTGACGGAGGAGTTTCTGAAGGAGAATCCCAACATTT
GCGCCTACATGGAACCGTCGCTGGACGCCAGGCAGGACATAGTAGTGGTG
GAGGTTCCGAAGCTGGGGAAAGAAGCGGCGACCAAGGCGATAAAGGAATG
GGGTCAACCAAAGTCAAAGATCACGCACCTCGTCTTCTGCACCACTTCCG
GCGTCGACATGCCCGGAGCTGATTATCAGCTCACCAAGCTTCTCGGCCTG
AGACCCTCCGTCAAGCGACTTATGATGTATCAACAGGGTTGCTTCGCCGG
CGGCACCGTCCTCCGCCTCGCCAAGGACCTCGCCGAGAACAACAAGGGCG
CTAGGGTTCTCGTCGTCTGCTCCGAGATCACCGCCGTCACGTTCCGCGGC
CCCTCCGAGACCCACCTCGACTCGCTCGTCGGCCAGGCGCTCTTCGGCGA
CGGCGCCGCCGCGGTGATTGTCGGCGCGGACCCCGACTTGGCCGTGGAGC
GCCCGCTCTTCGAGCTGGTGTCGGCGGCGCAGACGATTCTCCCCGACTCC
GACGGCGCCATCGACGGGCACCTCCGCGAGGTGGGGCTGACGTTCCACCT
TCTGAAGGACGTTCCAGGAATCATATCGAAGCACATCCAGAAGAGTCTGG

CGGACGCTTTTGCCCCCATTGGGATAAACGATTGGAACTCGCTGTTCTGG
ATCGCGCACCCGGGTGGACCGGCGATTCTGGACCAGGTGGAGGAGAAGCT
CCGGCTCAAACCGGAGAAGCTCCGGTCCACGCGGCACGTGCTGAGCGAGT
ACGGAAACATGTCGAGCGCGTGCGTGTTGTTCATTCTGGACGAGGTGCGG
AAGAGATCCAAAGAGGAAAAGAAGGGCACCACAGGGGAAGGGTTAGAGTG
GGGCGTGTTATTCGGGTTCGGGCCGGGTCTCACCGTCGAAACCGTTGTTC
TTCATAGCGTTCCCTTGGAGGGTTGA

1.3 DAC

DAC1 (C.cajan_42050)
ATGGGGGAATTCAAACACTTTGTAATTGTTAAGTTCAAGGAAGGTGTGGC
GGTTGAAGAGCTCACAAAAGGGATGGAGAAGTTCGTTTCAGAGATTGATG
CTGTCAAGTCCTTTGAATGGGGACAGGACATTGAAAGCCTAGATGTGTTG
AGACAAGGTTTCACTCATGCCTTCTTGATGACATTCGACAAGAAAGAAGA
CTTCGTTGCATTTCAGAGCCACCCCAATCATGTCGAGTTCTCCACAAAAT
TTTCAGCAGCTATAGAGAATATAGTGCTGCTGGATTTCCCATCTACTCTT
GTCAAGGCCCCAGCATGA

DAC2 (C.cajan_21648)
ATGGGGACTTTCAATCACTATGTGGTTGTTAAATTCAAGGATGGTGTGGC
AGTTGAAGAACTCATTCAAGGGTTGGAGAAAATGGTCTCTGGGATTGACC
ATGTCAAGTCCTTTGAATGGGGGAAGGACATCGAGAGTCATGAAATGCTG
AGACAAGGTTTCACTCATGCTTTCTTGATGACATTCAATGGGAAAGAGGA
GTGCAACGCATTTCAGACTCATCCAAATCATGTTGAGTTCTCCGGGATAT
TTTCACCTGCTATTGAGAAGATTGTGGTGCTGGATTTTCCATCTAAGCTT
GTCAAAGCACCAGCATGA

DAC3 (C.cajan_11068)
ATGGAAGAAGCAAAGGGATTAGTGAACCACGTAGTCCTCACAAAGTTCAA
GGACGACGTAGCTCCAGAGCGGATTGAAGACTTCATCAAGGGCTTTGCCA
ATCTAGTCAATCTCATCCCACCAATGAAATCATTCTCTTGGGGTAAGGAT
GTAAGTGACACAAGTATGCATCAAGGCTTCACTCATGTCTTTGTATCAAC
CTTTGAGAGTACAGAAGGCGTTGCAGAGTATTTAGCTCATCCAGCCCATG
TTGAATTTGCAAATTTGTTCCTTCTCTGTTTCGAGAAAGTCATTGCCATT
GACTATAAGCCTACCATTGTTAAATTGTAG

DAC4 (C.cajan_11069)
ATGGAGGAAGCAAAGGGATTAGTGAAACATATAGTCCTCGCAAAGTTTAA
GGACGACGTGGCTCCGGAGCGGATTGAAGACCTCATCAAGGCCTATGCTA
AACTAGTCCATCTCATCCCACCCATGAAGTCATTGCACTGGGGCAAGGAT
GTAAGTGACACAAATCTGCATCAAGGCTTCACTCATGTCTTTGAATCAAC
CTTTGAGAGTATAGAAGGTGTTGCAGAGTATGAAGCTCATCCTGCCCATG
TTGAATATGCAAATTTATTTCTTTCCTTTGCGGAGAAAATTGTTGTGGTT
AACTACAAGCCTACTACTGTAAACCTTTGA

DAC5 (C.cajan_11074)
ATGGAGGAAGCAAAAGGATTGGTGAAACACGTACTCGTCGCAAAGTTCAA
AGACGACATCACTCCAGAAAGAATTGAGGAGCTTATCAAGGACTATGCTA
ACCTAGTCAATCTCATCCCACCCATGAAGTCATTCCATTGGGGCAAGGAT
TTAAGTGCTGAAAACATGCATCAAGGTTTCACTCATGTCTTTGAATCAAC
CTTTGAGAGTACAGAAGGCGTTGCAGAGTATGTAGCTCATCCTGCCCATG
TTGAATATGCAAATTTGTTCCTTCCCTGTTTGGAGAAAGTCATTGTCATT
GACTATAAGCCTACCGTTGTTAAATTGTAA

1.4 BEC

BEC1 (C.cajan_31770)
ATGGAGGCCCTCACTTTCACTGAGGAGTTTACTAGCACTGTCCAGGCAGG
GAGGTTGTTCAAGGCTTTGATTCTTGATGCTCCTAACCTCATCCCCAAGC
TGATGCCTGAAGCCATTAAGAATGTTCAACTGATTGAAGGGAATGGTGGC
CCTGGAAGCATTCAGGAGATTACCATTGCTGAAGGTGATAAAATCAAAAG
CTTGAAGCATAGAATTGATGCAATAGACCAAGAGAAATTGACATACAGTT
ACGCAGTTATTGAGGGGGATGCTGCACTGGAAAAGGTGGACGCAATTTCA
CATGAGATTAAGTTCGAGGCCACTGAAAAGGGAGGTTGCAAGACTAAGAA
TGTGAGCAAGTATCAGCCCAAACCAGGGGTTGATGTCAAAGAGGAGGATT
TCAAGGCTGCAAGGGAGGAAGGCTTGGCTCTCCTCAAGGTGGTGGATGCC
TATCTTGTGGCAAACCCAGAGGCCTATGCTTGA

BEC2 (C.cajan_47832)
ATGGGTGTTTTCAGATTCGAGGATGAAACCACTTCTCCAGTGGCTCCAGC
TAAGCTTTACAAAGCCCTAGTTAAAGATGCTGATACCCTTGTCCCAAAGG
TTGTTGATGCCGTCAAGAGTGTAGAAATCGTTGAAGGAAATGGTGGCCCC
GGATCCATCAAGAAGCTCACTGTAGTTGAGGGTGGGGAGACGAAGCATGT
GTTGCACAAAGTTGAAACAATCGATGAGGCGAACTTGGGATACAGCTACA
GCATAGTTGGGGGTGCTGCGTTGCCAGACACAGTGGAGAAGATCACTTTT
GAGACAAAATTTGTTGCTGGACCCGATGGAGGGTCTATTGCCAAGCTGGC
TGCCACGCACCACACCAAAGGAGATGCCAAGCCCAGTGAGGAGGAGCTCA
AGATTGGGAAAGCCAAGGGAAATGCTTTCTTCAAAGCCATTGAGGCTTAC
CTTTTGGCCAATCCTGATTACTAG

BEC3 (C.cajan_42730)
ATGGGTGTTTTCTCATTTGAGGATGAAACCACATCACCTGTGGCTCCTGC
TAAGCTTTACAAAGCTCTAGTTAATGATGCTGATAACCTCATCCCAAAGG
TTGTTGAGGCTGTCAAGAGTGTTGAAATCGTTGAGGGAAATGGTGGCCCT
GGAACCATCAAGAAGCTCACTATAGTCGAGGAGGGGGAGACGAAGCATGT
GTTGCACAAAGTTGAAGGCATTGATGAGGCAAACTTTGGATACAGCTACA
GCATAGTTGGGGGTGTGGCTTTGCCAGACACAGTGGAGAAGATCACTTTT
GAGTCCAAATTCATTGCTGGCCCCAATGGAGGCTCTGTTGCCAAGCTCAG
TGTCAAGTTCCAAACCAAAGGAGATGCCAAGCCCAATGAAGAGGAACTCA
AAACTGGCAAAGCCAAGGGAAATGCTCTTTTCAAGGCCATTGAGGGTTAC
CTTTTGGCCAATCCTCATTACTGA

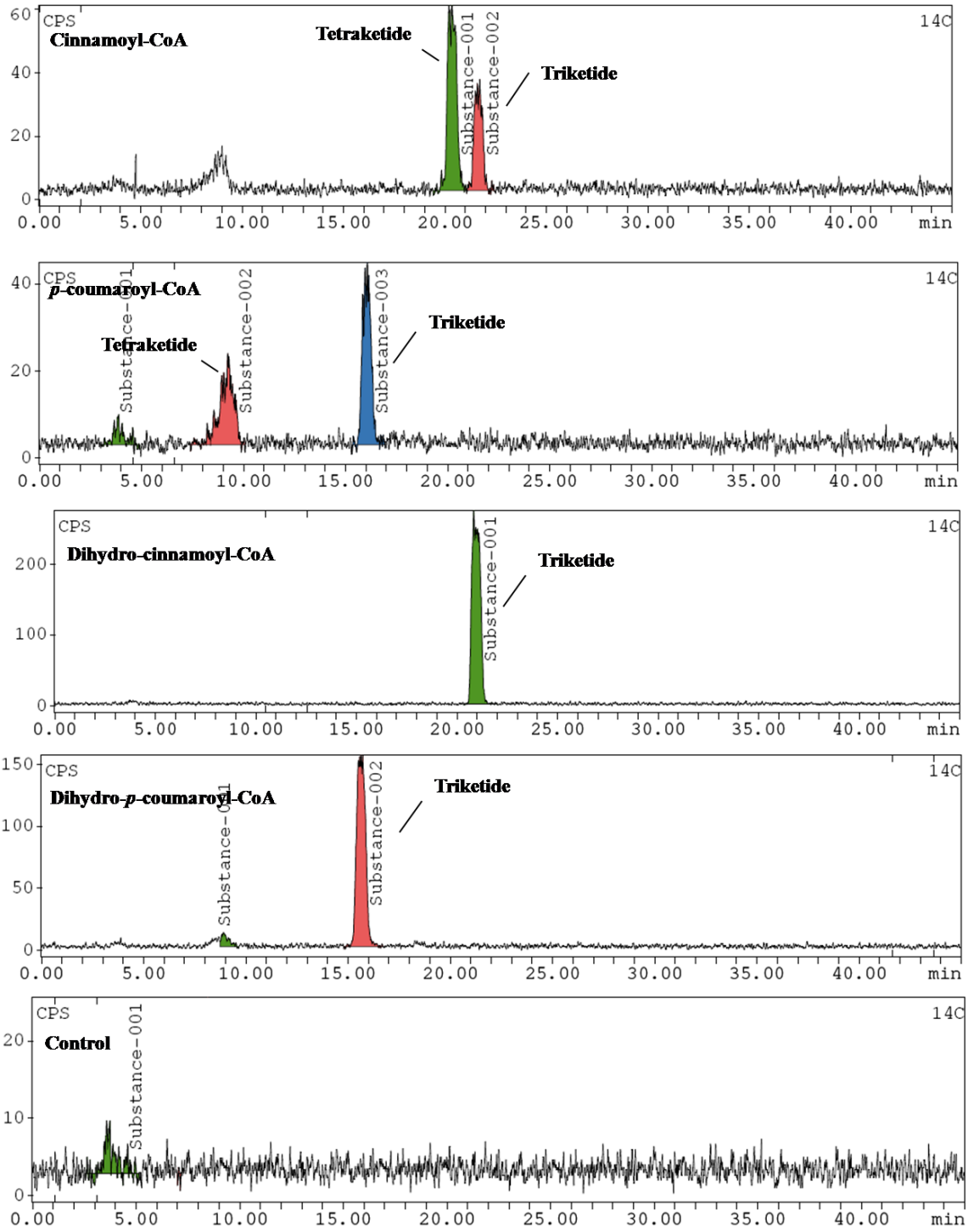
BEC4 (C.cajan_04110)
ATGGGTGTTTTCACATTCGAGGATGAAACCACTTCCCCAGTGGCTCCTGC
TAAGCTTTACAAAGCCTTTGTTACAGACGCCGATAACCTCATCCCAAAGG
TTGTTAGTGCCATCCAGGGTGTCGAAATCGTTGAAGGAAACGGTGGCCCC
GGAACCATCAAGAAGCTCACTGTAGTTGAGGATGGGGAGACTAAGCATGT
GTTGCACAAAATTGAATCAAGCGATGAAGCGAACTTTGCATACAGCTACA
GCGTAGTTGGGGGTGTGGCGTTGCCAGAGACAGCAGAGAAGATAACGTTT
TCGTCAAAATTTGTTGCTGGACCCAATGGAGGGTCTATTGCCAAGCTCAC
TGTGACGCACCACACCAAAGGAGATGCCAAGCCTGATGAAGAGGAACTCA
AGATTGGGAAAGCCAAGGGAAATGCTTTTTTCAAGGCCGTTGAGGCTTAC
CTTTTGGCCAATCCTGATTACTGA

BEC5 (C.cajan_40987)
ATGGGTGTTTTCACTTGTGAAGTCGAGCACGTTTCCCCCATCTCCGCTGC
AAAATTTTACAAAGCCGTTGTTGAAGATTGTGTCACTCTGTGGCCAAAAG
CAATGCCAAAAATCATTAAGAGCGTTGAATTCACTGAAGGAGATGGAGGG
CCAGGAACCATTTGCAAGCTTACTCTTCTTAAAGGATATTGGATGAGCCA
AGTAGACGCGATTGATAAAGCAAAGTTTGAGTTCAAGTATACTGTGATTG
AAGGAATGATATTGGGGAACCAATTGGAGAAGATCTATAATGATCTCAAA
GTGGTTGAAAGGGAAGATGGAGGATGCATTGTGAAGAAGATAATAAAATT
CTATACCAAAGACAATGAAATATACGATGAAAAAAATTTGAAGTTTAATA
AGGACAAAACACTTCTCTTCTTCAAGACCCTTGAGGATTTCCTTCTGGCT
AATCCTGATTACAACTAA

1.5 CHC

CHC1 (C.cajan_09652)
ATGGCATTCATGCGCTTCCCTTTCTCGTTTTCTCCGCTCCGACCGCCTAA
ATTTCCTCGCCCCTTCACTGCATTTGCTGTAGCCGCTGCCGCGGTCGGAG
CCTCTGCCGCGGTCGTTGCCGTCTCCTCCTCCGATCGCTCCTCTCTCTGG
AATGCGATGAGTTCCGTGGTGTCTTCTCAGCATTCATCTGCTCTCTGGGG
TTCACTCAGTTTGGAGGACAATGGTGTTCACGTGCTTGAGCGAAAGACCG
GAGTCTCCTTCCCTTCCGCCCTAGGCCCTTCTCTCAAGCTCTGTGGAATT
GGCTTACGCAAAAAAAGCGTTTTTGGATTGAAGAACATTGATGTTTACGC
ATTTGGTGTTTATGCTAATGATGAGGACATTAAGAAATATCTTTCTGAGA
AATATGGGAAGTTGCCTGTGTCGGAATTGCAGGGTAATAAGGAGTTTATT
GATGATCTTATGGAAAGTGATATACCCATGACGGTTAGACTTCAAATTGT
CTATGGCAGATTGAGCATTGGTTCCGTGCGTAGTGCATTTGAAGAGTCTG
TGGGAAGCAGACTGCGGAGGTTTGGGGGATCGGACAACAAAGAATTGCTT
CAAAGGTTCACTTCCCAGTTTAGAGATGAATTCAAAATACCTCGTGGATC
TGTAATCCATCTCTCAAAAGAAAAGGGCCACGTTCTTTGCACTTCAATTG
ATGGACAGGAAGTGGGATGCATCCAGAGCAAACTCCTATGTAAGTCAATT
CTAGATTTATATGTTGGAGAGGAAGCATTTGATGAACAGGCTAAAGAAGA
AATTGAGAACAACGTAGCTACATACCTCTAG

2. Chromatograms



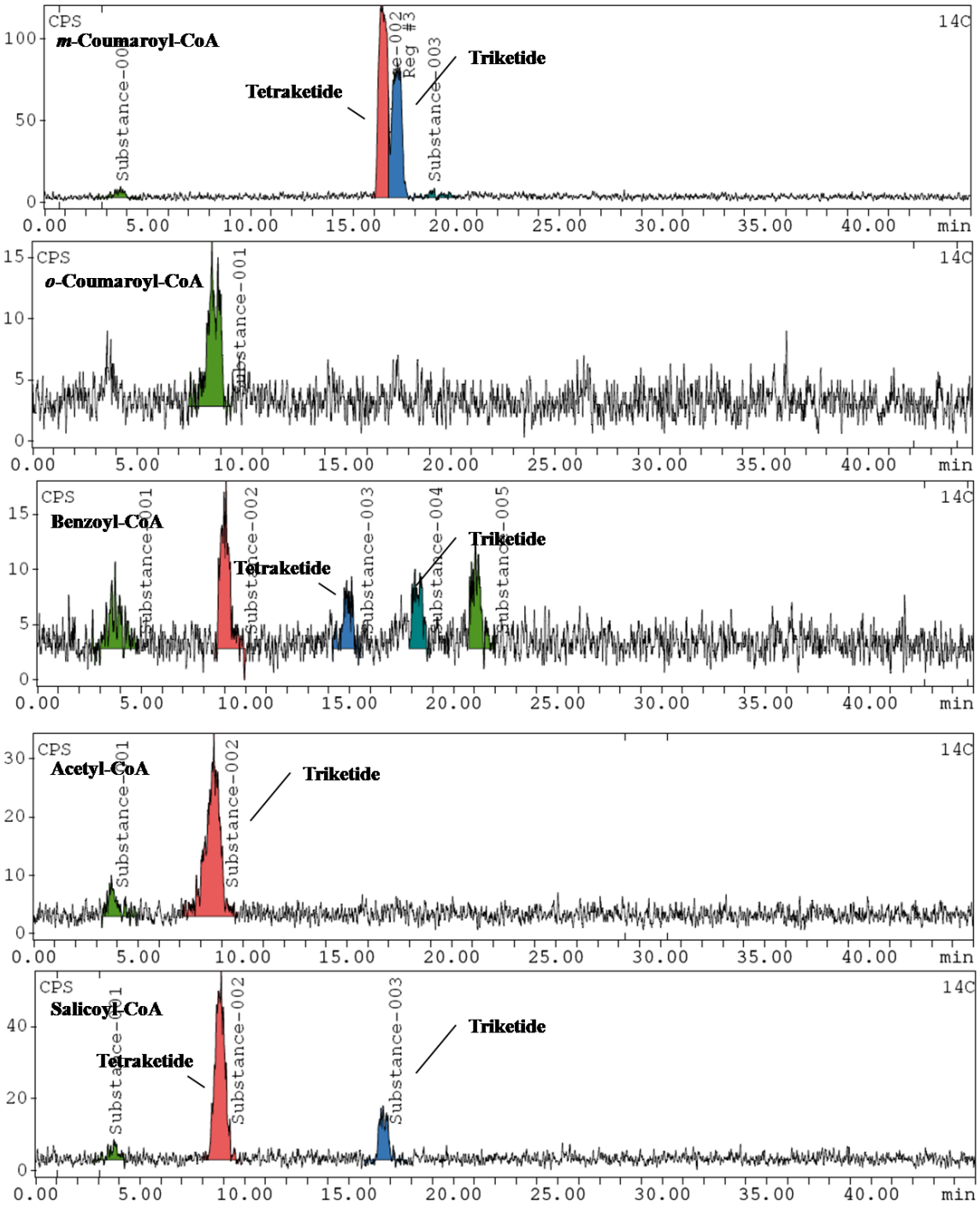
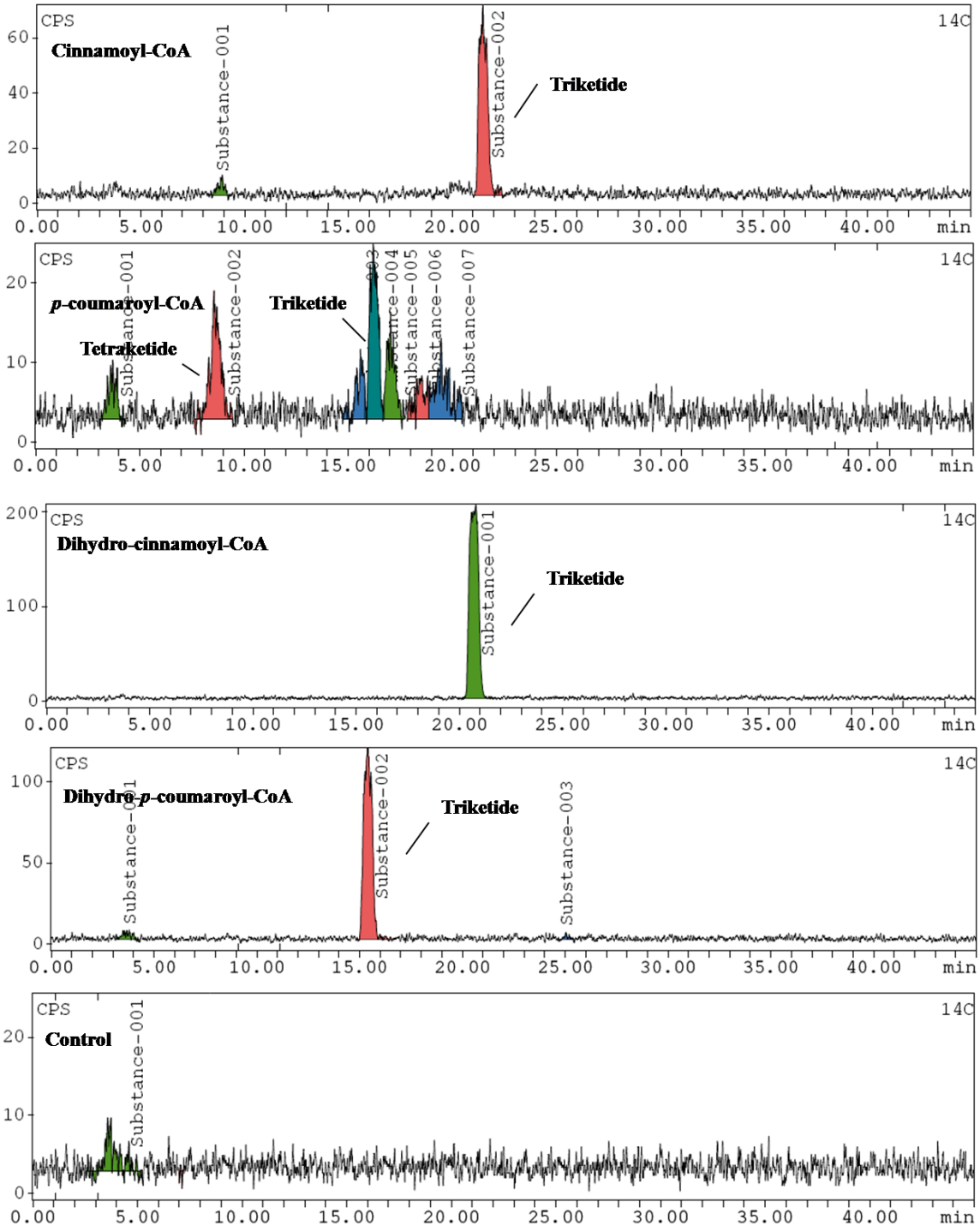


Fig. A1. Radiodetector-coupled HPLC analysis of PAS1 assays with different substrates, incubation with denatured protein was used as control.


CPS
14C
Cinnamoyl-CoA
Substance-001
Substance-002
Triketide
min
CPS
14C
p-coumaroyl-CoA
Tetraketide
Triketide
Substance-001
Substance-002
Substance-003
Substance-004
Substance-005
Substance-006
Substance-007
min
CPS
14C
Dihydro-cinnamoyl-CoA
Substance-001
Triketide
min
CPS
14C
Dihydro-p-coumaroyl-CoA
Substance-001
Substance-002
Triketide
Substance-003
min
CPS
14C
Control
Substance-001
min

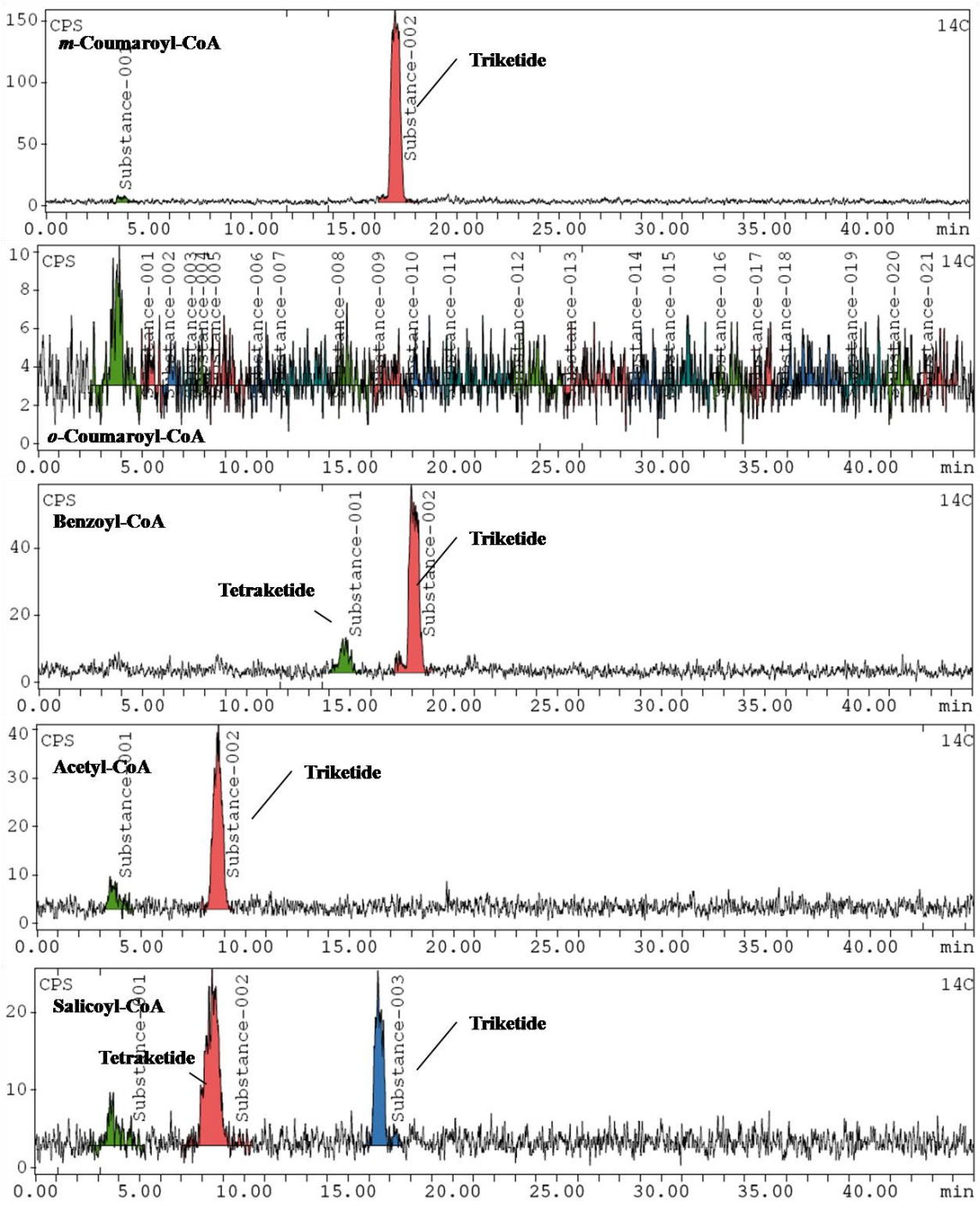


Fig. A2. Radiodetector-coupled HPLC analysis of PAS3 assays with different substrates, incubation with denatured protein was used as control.